



wwPDB EM Validation Summary Report ⓘ

Sep 16, 2021 – 11:02 pm BST

PDB ID : 7OQB
EMDB ID : EMD-13028
Title : The U2 part of *Saccharomyces cerevisiae* spliceosomal pre-A complex (delta BS-A ACT1)
Authors : Zhang, Z.; Rigo, N.; Dybkov, O.; Fourmann, J.; Will, C.L.; Kumar, V.; Urlaub, H.; Stark, H.; Luehrmann, R.
Deposited on : 2021-06-03
Resolution : 9.00 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

EMDB validation analysis : **FAILED**
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.23.1

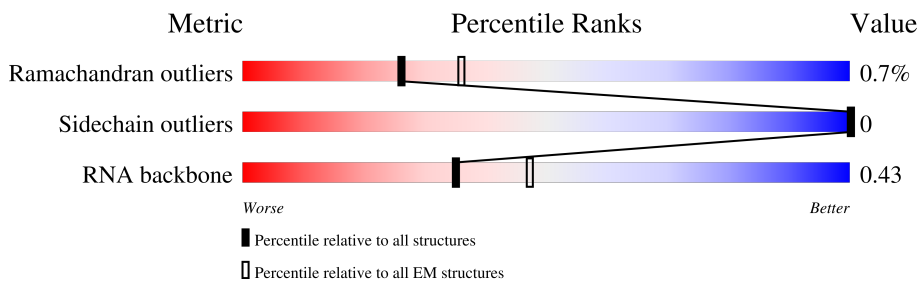
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 9.00 Å.

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



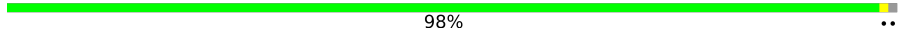



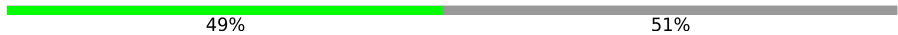




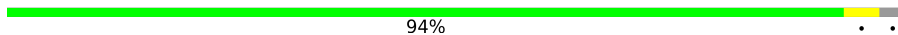


Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	O	971	80% 16%
2	I	318	93%
3	U	282	66% 33%
4	V	280	36% 63%
5	T	530	87% 13%
6	S	107	86% 14%
7	Q	436	50% 50%
8	P	1361	86% 13%
9	R	213	81% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	Z	84	 98%
11	W	238	 66% 5% 29%
12	Y	111	 75% 24%
13	s	196	 33% 67%
14	t	146	 49% 51%
15	u	110	 84% 16%
16	v	101	 81% 19%
17	w	93	 83% 17%
18	x	86	 85% 15%
19	y	77	 94%
20	p	849	 52% 48%
21	2	1175	 6% 88%

2 Entry composition i

There are 21 unique types of molecules in this entry. The entry contains 26426 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 U2 snRNP component HSH155.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
1	O	812	4108	2484	812	812	0	0

- Molecule 2 is a RNA chain called ACT1 pre-mRNA (delta-BS-A).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	I	23	480	216	78	163	23	0	0

- Molecule 3 is a protein called Pre-mRNA-splicing factor PRP11.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
3	U	188	943	567	188	188	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
U	121	UNK	-	insertion	UNP Q07350
U	122	UNK	-	insertion	UNP Q07350
U	123	UNK	-	insertion	UNP Q07350
U	124	UNK	-	insertion	UNP Q07350
U	125	UNK	-	insertion	UNP Q07350
U	126	UNK	-	insertion	UNP Q07350
U	127	UNK	-	insertion	UNP Q07350
U	128	UNK	-	insertion	UNP Q07350
U	129	UNK	-	insertion	UNP Q07350
U	130	UNK	-	insertion	UNP Q07350
U	131	UNK	-	insertion	UNP Q07350
U	132	UNK	-	insertion	UNP Q07350
U	133	UNK	-	insertion	UNP Q07350
U	134	UNK	-	insertion	UNP Q07350
U	135	UNK	-	insertion	UNP Q07350

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
U	136	UNK	-	insertion	UNP Q07350

- Molecule 4 is a protein called Pre-mRNA-splicing factor PRP21.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	V	103	515	309	103	103	0	0

- Molecule 5 is a protein called Pre-mRNA-splicing factor PRP9.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	T	462	2318	1394	462	462	0	0

- Molecule 6 is a protein called Pre-mRNA-splicing factor RDS3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
6	S	92	460	276	92	92	0	0

- Molecule 7 is a protein called Cold sensitive U2 snRNA suppressor 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
7	Q	220	1122	682	220	220	0	0

- Molecule 8 is a protein called Pre-mRNA-splicing factor RSE1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
8	P	1186	5972	3600	1186	1186	0	0

- Molecule 9 is a protein called Protein HSH49.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
9	R	173	868	522	173	173	0	0

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

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
10	Z	83	412	246	83	83	0	0

- Molecule 11 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
11	W	170	862	522	170	170	0	0

- Molecule 12 is a protein called U2 small nuclear ribonucleoprotein B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
12	Y	84	418	250	84	84	0	0

- Molecule 13 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
13	s	65	323	193	65	65	0	0

- Molecule 14 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
14	t	72	363	219	72	72	0	0

- Molecule 15 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
15	u	92	463	279	92	92	0	0

- Molecule 16 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
16	v	82	412	248	82	82	0	0

- Molecule 17 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
17	w	77	389	235	77	77	0	0

- Molecule 18 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
18	x	73	365	219	73	73	0	0

- Molecule 19 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
19	y	75	373	223	75	75	0	0

- Molecule 20 is a protein called Pre-mRNA-processing ATP-dependent RNA helicase PRP5.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
20	p	444	2239	1351	444	444	5	0

- Molecule 21 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
21	2	143	3021	1351	511	1017	142	0	0

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	160894	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 ($e^-/\text{\AA}^2$)	44	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	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 >5	RMSZ	# Z >5
1	O	0.42	0/4149	0.77	30/5819 (0.5%)
2	I	0.69	6/534 (1.1%)	0.86	0/827
3	U	0.22	0/867	0.43	0/1208
4	V	0.38	0/515	0.43	0/719
5	T	0.27	0/2324	0.44	0/3248
6	S	0.27	0/463	0.49	0/645
7	Q	0.27	0/1137	0.47	0/1593
8	P	0.28	1/6009 (0.0%)	0.54	0/8407
9	R	0.28	0/869	0.46	0/1209
10	Z	0.26	0/412	0.41	0/573
11	W	0.32	0/869	0.60	0/1219
12	Y	0.27	0/418	0.49	0/582
13	s	0.30	0/322	0.57	0/446
14	t	0.33	0/364	0.56	0/507
15	u	0.32	0/465	0.53	0/650
16	v	0.29	0/415	0.54	0/579
17	w	0.29	0/392	0.54	0/546
18	x	0.31	0/367	0.58	0/510
19	y	0.26	0/374	0.50	0/520
20	p	0.55	1/2269 (0.0%)	0.66	3/3172 (0.1%)
21	2	4.64	44/3363 (1.3%)	2.45	107/5218 (2.1%)
All	All	1.68	52/26897 (0.2%)	1.06	140/38197 (0.4%)

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
5	T	0	1
8	P	0	2
11	W	0	1
21	2	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	0	6

The worst 5 of 52 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
21	2	35	U	C1'-N1	151.34	3.75	1.48
21	2	42	U	C1'-N1	150.88	3.75	1.48
21	2	44	U	C1'-N1	149.94	3.73	1.48
20	p	271	THR	C-N	20.19	1.80	1.34
21	2	1161	U	O3'-P	-15.60	1.42	1.61

The worst 5 of 140 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
21	2	44	U	C6-N1-C1'	-74.41	17.03	121.20
21	2	42	U	C6-N1-C1'	-73.72	17.98	121.20
21	2	35	U	C6-N1-C1'	-73.61	18.15	121.20
21	2	44	U	O4'-C1'-N1	-27.52	86.18	108.20
21	2	42	U	O4'-C1'-N1	-21.00	91.40	108.20

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
21	2	141	A	Sidechain
8	P	1013	ASP	Peptide
8	P	1014	LYS	Peptide
5	T	458	SER	Peptide
11	W	16	VAL	Peptide

5.2 Too-close contacts [i](#)

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

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

entries.

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	O	810/971 (83%)	770 (95%)	37 (5%)	3 (0%)	34	72
3	U	166/282 (59%)	141 (85%)	24 (14%)	1 (1%)	25	66
4	V	101/280 (36%)	90 (89%)	10 (10%)	1 (1%)	15	55
5	T	454/530 (86%)	414 (91%)	40 (9%)	0	100	100
6	S	90/107 (84%)	79 (88%)	11 (12%)	0	100	100
7	Q	214/436 (49%)	202 (94%)	11 (5%)	1 (0%)	29	69
8	P	1170/1361 (86%)	1059 (90%)	104 (9%)	7 (1%)	25	66
9	R	165/213 (78%)	161 (98%)	3 (2%)	1 (1%)	25	66
10	Z	81/84 (96%)	76 (94%)	4 (5%)	1 (1%)	13	50
11	W	168/238 (71%)	129 (77%)	28 (17%)	11 (6%)	1	16
12	Y	82/111 (74%)	76 (93%)	5 (6%)	1 (1%)	13	50
13	s	61/196 (31%)	58 (95%)	3 (5%)	0	100	100
14	t	68/146 (47%)	67 (98%)	1 (2%)	0	100	100
15	u	90/110 (82%)	89 (99%)	1 (1%)	0	100	100
16	v	80/101 (79%)	77 (96%)	3 (4%)	0	100	100
17	w	73/93 (78%)	72 (99%)	1 (1%)	0	100	100
18	x	71/86 (83%)	69 (97%)	2 (3%)	0	100	100
19	y	73/77 (95%)	64 (88%)	6 (8%)	3 (4%)	3	23
20	p	445/849 (52%)	431 (97%)	13 (3%)	1 (0%)	47	81
All	All	4462/6271 (71%)	4124 (92%)	307 (7%)	31 (1%)	26	63

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
7	Q	368	ILE
8	P	1299	ILE
11	W	34	LEU
11	W	52	LYS
19	y	50	ASP

5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	O	42/867 (5%)	42 (100%)	0	100	100
3	U	7/236 (3%)	7 (100%)	0	100	100
4	V	1/259 (0%)	1 (100%)	0	100	100
5	T	10/492 (2%)	10 (100%)	0	100	100
6	S	4/97 (4%)	4 (100%)	0	100	100
7	Q	18/392 (5%)	18 (100%)	0	100	100
8	P	45/1244 (4%)	45 (100%)	0	100	100
9	R	5/189 (3%)	5 (100%)	0	100	100
10	Z	1/76 (1%)	1 (100%)	0	100	100
11	W	8/219 (4%)	8 (100%)	0	100	100
12	Y	1/100 (1%)	1 (100%)	0	100	100
13	s	1/176 (1%)	1 (100%)	0	100	100
14	t	3/129 (2%)	3 (100%)	0	100	100
15	u	3/103 (3%)	3 (100%)	0	100	100
16	v	4/89 (4%)	4 (100%)	0	100	100
17	w	5/82 (6%)	5 (100%)	0	100	100
18	x	3/77 (4%)	3 (100%)	0	100	100
19	y	2/66 (3%)	2 (100%)	0	100	100
20	p	17/768 (2%)	17 (100%)	0	100	100
All	All	180/5661 (3%)	180 (100%)	0	100	100

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

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

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	I	22/318 (6%)	11 (50%)	0
21	2	138/1175 (11%)	53 (38%)	27 (19%)
All	All	160/1493 (10%)	64 (40%)	27 (16%)

5 of 64 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	I	247	U
2	I	248	A
2	I	249	C
2	I	250	U
2	I	251	A

5 of 27 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
21	2	1119	C
21	2	1122	U
21	2	1144	U
21	2	1121	U
21	2	1123	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 monosaccharides 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

There are no chain breaks in this entry.