



wwPDB EM Validation Summary Report ⓘ

Mar 17, 2024 – 12:08 AM JST

PDB ID : 8Y7E
EMDB ID : EMD-39013
Title : Cryo-EM Structure of the human minor pre-B complex (pre-precatalytic spliceosome) U12 snRNP part
Authors : Bai, R.; Yuan, M.; Zhang, P.; Luo, T.; Shi, Y.; Wan, R.
Deposited on : 2024-02-04
Resolution : 4.66 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

EMDB validation analysis : **FAILED**
Mogul : 1.8.5 (274361), CSD as541be (2020)
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

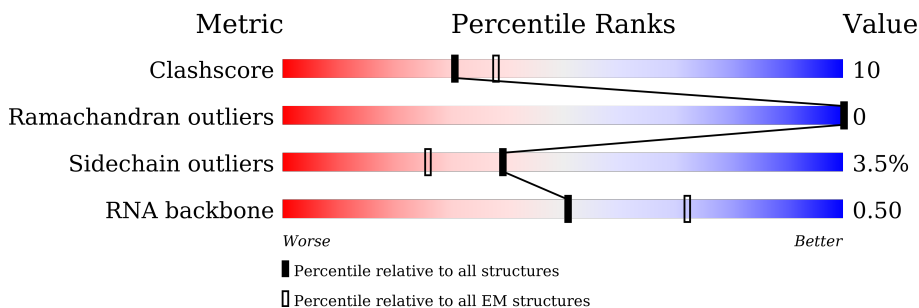
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.66 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	158937	4297
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	A	280	5% . . 89%
2	1	1304	62% 11% . 26%
3	2	895	21% . 76%
4	3	1217	77% 19% ..
5	4	424	14% . 82%
6	5	125	71% 15% . 13%
7	6	110	76% 18% . 5%
8	7	86	86% 8% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	H	150	
10	h	126	
11	i	240	
12	j	119	
13	k	118	
14	l	92	
15	m	86	
16	n	76	
17	v	230	

2 Entry composition [i](#)

There are 18 unique types of molecules in this entry. The entry contains 26949 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 pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	A	30	613	276	86	221	30	0	0

- Molecule 2 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	N	O	P	S		
2	1	971	7766	4961	1342	1415	1	47	0	0

- Molecule 3 is a protein called Splicing factor 3B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	2	216	1674	1067	296	305	6	0	0

- Molecule 4 is a protein called Splicing factor 3B subunit 3.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	N	O	P	S		
4	3	1193	9352	5932	1590	1784	1	45	0	0

- Molecule 5 is a protein called Splicing factor 3B subunit 4.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	4	78	383	227	78	78	0	0

- Molecule 6 is a protein called Splicing factor 3B subunit 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	5	109	906	582	157	163	4	0	0

- Molecule 7 is a protein called PHD finger-like domain-containing protein 5A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	6	105	811	502	145	151	13	0	0

- Molecule 8 is a protein called Splicing factor 3B subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	7	81	669	422	117	124	6	0	0

- Molecule 9 is a RNA chain called U12 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
9	H	50	1079	482	203	344	50	0	0

- Molecule 10 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
10	h	80	393	233	80	80	0	0

- Molecule 11 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
11	i	86	422	250	86	86	0	0

- Molecule 12 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
12	j	82	406	242	82	82	0	0

- Molecule 13 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
13	k	85	422	252	85	85	0	0

- Molecule 14 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms				AltConf	Trace
14	l	79	Total	C	N	O	0	0
			391	233	79	79		

- Molecule 15 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	m	74	Total	C	N	O	0	0
			361	213	74	74		

- Molecule 16 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms				AltConf	Trace
16	n	68	Total	C	N	O	0	0
			334	198	68	68		

- Molecule 17 is a protein called Sodium channel modifier 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	v	119	Total	C	N	O	S	0	0
			963	601	179	178	5		

- Molecule 18 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
18	6	3	Total	Zn	0
			3	3	
18	v	1	Total	Zn	0
			1	1	

PRO
PRO
VAL
PRO
PRO
ARG
GLY
PRO
LEU
LEU
LEU
PRO
GLN

• Molecule 6: Splicing factor 3B subunit 6



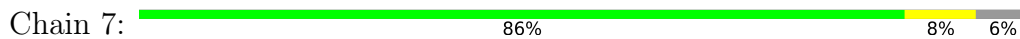
MET
ALA
MET
GLN
ALA
ALA
LYS
ARG
ALA
ASN
ARG
ILE
R12
R24
P27
M85
F59
Q47
I48
R49
V62
A73
H76
L77
F80
N81
V82
C83
N84
R85
T86
L87
V88
V89
R96
T103
H120
THR
ASP
PRO
PRO
LYS

• Molecule 7: PHD finger-like domain-containing protein 5A



MET
A2
D7
L22
K25
K29
C30
D34
S35
L42
I45
N50
C58
V59
I60
G61
G62
D68
A69
R82
L91
K95
T96
Y100
E101
R102
G106
PHE
LYS
LYS
ARG

• Molecule 8: Splicing factor 3B subunit 5



M1
A24
L32
C41
D48
N51
E57
M58
D61
LYS
PRO
GLU
GLU
ASN

• Molecule 9: U12 snRNA



A
U
G
C
C
U
U
A
A
A
A
C
C
U
U
A14
A17
G18
G18
A29
C31
G32
G38
G39
G40
C44
G45
A50
A51
U
C
C
C
U
U
A
C
C
A
C
C
U
U
C
G
C
C
U
U
A
A
U
U
G
U
C
G
C
A
A
G
A
C
G73
A74
U77
U78
U79
U80

G81
A82
G83
C84
G
G
U
U
U
A
A
A
A
A
G
G
U
U
C
C
C
C
G

G
A
U
U
C
U
G
C
C
C
G

• Molecule 10: Small nuclear ribonucleoprotein Sm D3



MET
SER
ILE
G4
T79
ALA
PRO
M82
S85
MET
LYS
ASN
LYS
ASN
LYS
GLN
GLN
GLY
SER
GLY
ALA
ALA
ARG
ILE
ILE
GLY
GLY
ALA
ALA
ALA
ALA
ILE
ILE
LEU
LYS
LYS
ALA
ALA
ALA
ALA
GLY
VAL
ALA
VAL
PRO
PRO
MET
MET
PRO
GLN
ALA
PRO
PRO
GLY
LEU
ALA
ALA
GLY
ASN
PRO
ILE
VAL
PHE
ARG
GLN
LYS
VAL
GLY
GLY
GLY
GLY
PRO
SER
SER
GLN
GLN
VAL
MET
THR

• Molecule 11: Small nuclear ribonucleoprotein-associated proteins B and B'



MET
THR
VAL
G4
E89
THR
THR
GLY
ILE
ALA
ALA
ARG
VAL
PRO
LEU
ALA
GLY
ALA
ALA
ALA
GLY
GLY
PRO
ILE
ILE
GLY
ARG
ALA
ALA
ALA
GLY
ILE
LEU
LEU
PRO
ALA
ALA
VAL
VAL
PRO
MET
MET
PRO
GLN
ALA
PRO
PRO
GLY
LEU
ALA
ALA
GLY
ASN
PRO
ILE
VAL
PHE
ARG
GLN
LYS
VAL
GLY
GLY
GLY
GLY
PRO
SER
SER
GLN
GLN
VAL
MET
THR

S219	D2226
LEU	PRO
LEU	LEU
ASP	

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	107724	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 ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	81000	Depositor
Image detector	GATAN K3 (6k x 4k)	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, SEP, TPO

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	A	0.31	0/678	0.70	0/1048
2	1	1.14	0/7908	0.79	0/10698
3	2	1.01	0/1710	0.70	0/2306
4	3	1.02	0/9531	0.74	0/12931
5	4	0.60	0/382	0.70	0/529
6	5	0.88	0/925	0.72	0/1247
7	6	1.15	0/825	0.77	0/1106
8	7	1.29	0/688	0.85	0/930
9	H	0.40	0/1209	0.70	0/1882
10	h	0.48	0/391	0.63	0/540
11	i	0.51	0/421	0.73	0/583
12	j	0.57	0/405	0.73	0/563
13	k	0.67	0/420	0.81	0/583
14	l	0.64	0/390	0.81	0/542
15	m	0.75	0/360	0.81	0/497
16	n	0.54	0/332	0.72	0/458
17	v	0.94	0/984	0.66	0/1326
All	All	0.99	0/27559	0.75	0/37769

There are no bond length outliers.

There are no bond angle outliers.

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	613	0	315	24	0
2	1	7766	0	7934	140	0
3	2	1674	0	1580	23	0
4	3	9352	0	9273	202	0
5	4	383	0	173	25	0
6	5	906	0	913	25	0
7	6	811	0	788	17	0
8	7	669	0	631	6	0
9	H	1079	0	541	14	0
10	h	393	0	170	0	0
11	i	422	0	177	0	0
12	j	406	0	170	0	0
13	k	422	0	175	0	0
14	l	391	0	163	0	0
15	m	361	0	158	0	0
16	n	334	0	143	0	0
17	v	963	0	945	0	0
18	6	3	0	0	0	0
18	v	1	0	0	0	0
All	All	26949	0	24249	450	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 10.

The worst 5 of 450 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)
2:1:641:ILE:HD11	2:1:675:MET:CE	1.58	1.31
5:4:14:THR:HA	5:4:59:VAL:O	1.32	1.23
5:4:12:ASP:CB	5:4:67:ALA:HB1	1.80	1.11
2:1:641:ILE:CD1	2:1:675:MET:CE	2.27	1.10
4:3:950:ALA:HB2	4:3:989:TYR:OH	1.53	1.07

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	
2	1	964/1304 (74%)	948 (98%)	16 (2%)	0	100	100
3	2	206/895 (23%)	197 (96%)	9 (4%)	0	100	100
4	3	1184/1217 (97%)	1137 (96%)	47 (4%)	0	100	100
5	4	76/424 (18%)	71 (93%)	5 (7%)	0	100	100
6	5	107/125 (86%)	104 (97%)	3 (3%)	0	100	100
7	6	103/110 (94%)	98 (95%)	5 (5%)	0	100	100
8	7	79/86 (92%)	78 (99%)	1 (1%)	0	100	100
10	h	76/126 (60%)	74 (97%)	2 (3%)	0	100	100
11	i	84/240 (35%)	82 (98%)	2 (2%)	0	100	100
12	j	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
13	k	81/118 (69%)	78 (96%)	3 (4%)	0	100	100
14	l	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
15	m	72/86 (84%)	68 (94%)	4 (6%)	0	100	100
16	n	64/76 (84%)	62 (97%)	2 (3%)	0	100	100
17	v	115/230 (50%)	113 (98%)	2 (2%)	0	100	100
All	All	3368/5248 (64%)	3263 (97%)	105 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	
2	1	842/1103 (76%)	808 (96%)	34 (4%)	31	56
3	2	163/776 (21%)	158 (97%)	5 (3%)	40	63
4	3	1031/1050 (98%)	997 (97%)	34 (3%)	38	61
6	5	97/109 (89%)	96 (99%)	1 (1%)	76	86
7	6	90/95 (95%)	87 (97%)	3 (3%)	38	61
8	7	72/77 (94%)	72 (100%)	0	100	100
17	v	104/199 (52%)	98 (94%)	6 (6%)	20	47
All	All	2399/3409 (70%)	2316 (96%)	83 (4%)	39	60

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	3	612	ASN
4	3	1209	GLU
4	3	686	LEU
4	3	982	GLU
7	6	35	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	1	1002	ASN
2	1	1252	GLN
4	3	51	HIS
4	3	304	GLN
4	3	805	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	29/280 (10%)	13 (44%)	4 (13%)
9	H	48/150 (32%)	8 (16%)	0
All	All	77/430 (17%)	21 (27%)	4 (5%)

5 of 21 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	196	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	203	A
1	A	207	A
1	A	208	U
1	A	210	G

All (4) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	206	C
1	A	213	C
1	A	217	U
1	A	219	U

5.4 Non-standard residues in protein, DNA, RNA chains

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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
2	SEP	1	129	2	8,9,10	1.08	1 (12%)	8,12,14	1.72	3 (37%)
4	TPO	3	613	4	8,10,11	0.90	0	10,14,16	1.63	2 (20%)

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
2	SEP	1	129	2	-	4/5/8/10	-
4	TPO	3	613	4	-	3/9/11/13	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	1	129	SEP	P-O1P	2.15	1.57	1.50

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	3	613	TPO	P-OG1-CB	-4.43	109.81	123.21
2	1	129	SEP	P-OG-CB	-2.82	110.54	118.30
4	3	613	TPO	O-C-CA	-2.27	118.83	124.78
2	1	129	SEP	OG-CB-CA	2.19	110.28	108.14
2	1	129	SEP	O3P-P-OG	2.18	112.54	106.73

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	1	129	SEP	N-CA-CB-OG
2	1	129	SEP	CB-OG-P-O1P
2	1	129	SEP	CB-OG-P-O2P
2	1	129	SEP	CB-OG-P-O3P
4	3	613	TPO	O-C-CA-CB

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	1	129	SEP	2	0
4	3	613	TPO	1	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.