



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Aug 22, 2017 – 07:33 AM EDT

PDB ID : 5GAO
EMDB ID: : EMD-8013
Title : Head region of the yeast spliceosomal U4/U6.U5 tri-snRNP
Authors : Nguyen, T.H.D.; Galej, W.P.; Bai, X.C.; Oubridge, C.; Scheres, S.H.W.; Newman, A.J.; Nagai, K.
Deposited on : unknown
Resolution : 3.60 Å(reported)
Based on PDB ID : 4BGD

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
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

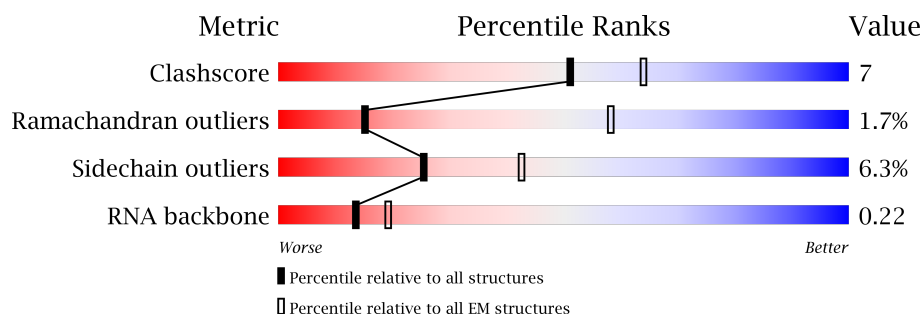
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.60 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 ≥ 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	k	196	39% 59%
2	l	146	60% 38%
3	m	110	80% 5% 15%
4	n	101	79% 19%
5	p	94	76% 5% 19%
6	q	86	83% 16%
7	r	77	86% 10%
8	E	338	98%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	B	2163	<div><div></div><div>59%</div><div>18%</div><div>•</div><div>21%</div></div>
10	V	96	<div><div></div><div>17%</div><div>35%</div><div>7%</div><div>41%</div></div>
11	A	267	<div><div></div><div>76%</div><div>17%</div><div>•</div><div>•</div></div>

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 23073 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 Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	k	80	Total	C	N	O	S	0	0
			635	406	115	111	3		

- Molecule 2 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	l	91	Total	C	N	O	S	0	0
			720	455	129	134	2		

- Molecule 3 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	m	94	Total	C	N	O	S	0	0
			737	474	140	119	4		

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

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

- Molecule 5 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	p	76	Total	C	N	O	S	0	0
			580	382	93	102	3		

- Molecule 6 is a protein called Small nuclear ribonucleoprotein F.

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

- Molecule 7 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	r	69	Total	C	N	O	S	0	0
			526	336	93	95	2		

- Molecule 8 is a protein called Snu66.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	E	338	Total	C	N	O		0	0
			1763	1063	352	348			

- Molecule 9 is a protein called Pre-mRNA-splicing helicase BRR2.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	B	1710	Total	C	N	O	S	1	0
			13690	8772	2283	2579	56		

- Molecule 10 is a RNA chain called Saccharomyces cerevisiae strain UOA_M2 chromosome 5 sequence.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	V	57	Total	C	N	O	P	0	0
			1209	542	212	398	57		

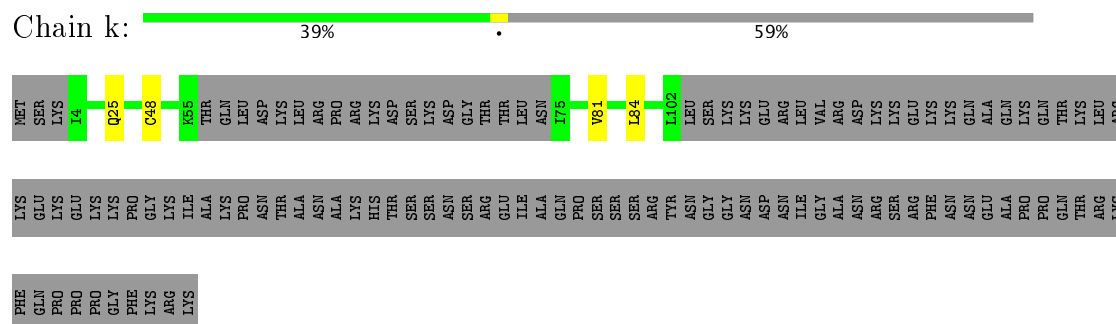
- Molecule 11 is a protein called Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	A	255	Total	C	N	O	S	0	0
			2015	1300	329	380	6		

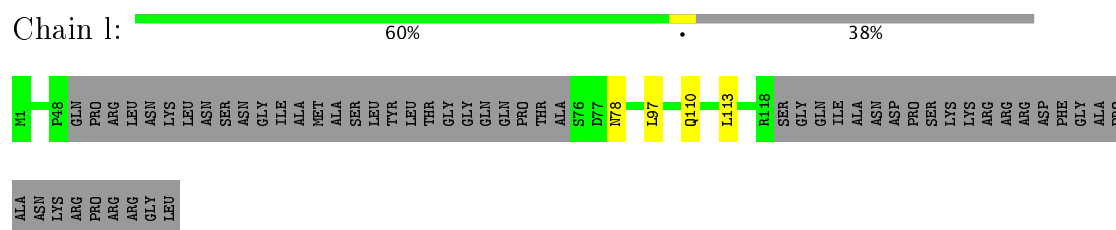
3 Residue-property plots [i](#)

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.

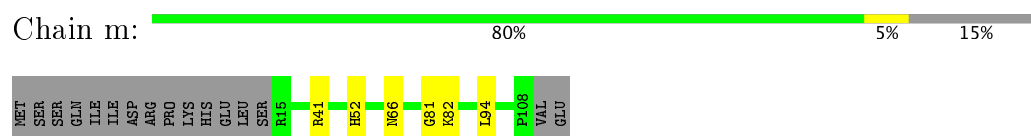
- Molecule 1: Small nuclear ribonucleoprotein-associated protein B



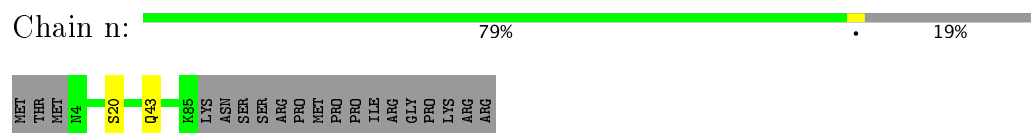
- Molecule 2: Small nuclear ribonucleoprotein Sm D1



- Molecule 3: Small nuclear ribonucleoprotein Sm D2



- Molecule 4: Small nuclear ribonucleoprotein Sm D3



- Molecule 5: Small nuclear ribonucleoprotein E





G2311	I2321	M2322	F2329	E2330	Q2338	L2339	L2340	F2349	G2354	M2355	V2356	Y2357	F2374	E2381	R2388	P2389	F2395	S2396	P2397	L2398	D2401	GLU	GLU	LEU	GLU	ALA	GLU	GLN	ILE	ASP	VAL	PHE	SER	G2147	M2152	R2153	F2154	S2155	M2159	L2162	V2163	L2164	R2165	L2166	K2167	M2168	L2169	Y2170	V2182	M2188	L2189	L2190	I2194	S2197	A2204	Y2208	T2223	V2224	V2227	P2228	Q2229	I2238	L2244	L2247	E2251	I2259	T2260	T2261	I2293	V2300	M2306	D2309	E2310
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	140155	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	38	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	35714	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	k	0.42	0/640	0.66	0/860
10	V	0.33	0/1350	0.77	0/2097
11	A	0.48	0/2064	0.67	0/2801
2	l	0.44	0/725	0.68	0/980
3	m	0.44	0/749	0.64	0/1009
4	n	0.37	0/634	0.57	0/859
5	p	0.48	0/590	0.61	0/802
6	q	0.42	0/585	0.63	0/791
7	r	0.38	0/529	0.57	0/711
8	E	0.41	0/1751	0.53	0/2411
9	B	0.46	0/13989	0.72	1/18966 (0.0%)
All	All	0.45	0/23606	0.69	1/32287 (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
1	k	0	1
9	B	0	5
All	All	0	6

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	B	643	ARG	NE-CZ-NH1	5.31	122.96	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
9	B	1200	PRO	Peptide
9	B	1484	TYR	Peptide
9	B	790	ASP	Peptide
9	B	791	PRO	Peptide
1	k	84	LEU	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	k	635	0	681	0	0
2	l	720	0	772	0	0
3	m	737	0	767	0	0
4	n	625	0	647	0	0
5	p	580	0	602	0	0
6	q	573	0	572	0	0
7	r	526	0	555	0	0
8	E	1763	0	1764	3	0
9	B	13690	0	13704	234	0
10	V	1209	0	612	15	0
11	A	2015	0	1970	28	0
All	All	23073	0	22646	267	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 267 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)
9:B:1700:ILE:HD11	9:B:1740:LEU:HD13	1.37	1.06
9:B:1208:ALA:HB2	9:B:1218:PHE:CD2	2.23	0.74
9:B:1220:ILE:HD13	9:B:1241:LEU:HD11	1.71	0.72
9:B:1183:ILE:CD1	9:B:1192:VAL:HG11	2.19	0.71
9:B:1224:ALA:HB3	9:B:1264:VAL:HA	1.72	0.71

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	k	76/196 (39%)	67 (88%)	8 (10%)	1 (1%)	14	57
2	l	87/146 (60%)	74 (85%)	12 (14%)	1 (1%)	17	61
3	m	92/110 (84%)	80 (87%)	10 (11%)	2 (2%)	8	47
4	n	80/101 (79%)	70 (88%)	10 (12%)	0	100	100
5	p	72/94 (77%)	64 (89%)	8 (11%)	0	100	100
6	q	70/86 (81%)	60 (86%)	10 (14%)	0	100	100
7	r	65/77 (84%)	58 (89%)	6 (9%)	1 (2%)	12	55
8	E	308/338 (91%)	266 (86%)	38 (12%)	4 (1%)	14	57
9	B	1707/2163 (79%)	1499 (88%)	174 (10%)	34 (2%)	9	49
11	A	253/267 (95%)	213 (84%)	34 (13%)	6 (2%)	7	45
All	All	2810/3578 (78%)	2451 (87%)	310 (11%)	49 (2%)	15	52

5 of 49 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
9	B	448	LEU
9	B	771	THR
9	B	792	SER
9	B	985	GLU
9	B	1370	SER

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	k	71/176 (40%)	69 (97%)	2 (3%)	49	80
2	l	85/129 (66%)	82 (96%)	3 (4%)	41	76
3	m	78/103 (76%)	74 (95%)	4 (5%)	28	66
4	n	69/89 (78%)	67 (97%)	2 (3%)	48	80
5	p	65/83 (78%)	60 (92%)	5 (8%)	15	52
6	q	63/77 (82%)	62 (98%)	1 (2%)	68	88
7	r	57/66 (86%)	55 (96%)	2 (4%)	41	76
8	E	20/20 (100%)	18 (90%)	2 (10%)	9	41
9	B	1536/1955 (79%)	1428 (93%)	108 (7%)	18	56
11	A	220/236 (93%)	206 (94%)	14 (6%)	20	59
All	All	2264/2934 (77%)	2121 (94%)	143 (6%)	25	60

5 of 143 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	B	1057	LEU
9	B	1253	TYR
11	A	2170	TYR
9	B	1085	SER
9	B	1137	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
9	B	804	HIS
9	B	1280	ASN
11	A	2237	GLN
9	B	739	GLN
9	B	2069	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
10	V	55/96 (57%)	34 (61%)	5 (9%)

5 of 34 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
10	V	74	U
10	V	75	U
10	V	78	A
10	V	80	A
10	V	81	C

All (5) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
10	V	80	A
10	V	83	A
10	V	88	G
10	V	143	A
10	V	150	G

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
8	E	14

The worst 5 of 14 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	132:ALA	C	150:ALA	N	37.33
1	E	170:ALA	C	180:ALA	N	34.57
1	E	102:ALA	C	120:ALA	N	31.59
1	E	299:ALA	C	310:ALA	N	27.52
1	E	226:ALA	C	240:ALA	N	26.36