W O R L D W I D E

wwPDB

PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report (i)

May 29, 2020 – 04:30 am BST

PDB ID	:	4BY9
Title	:	The structure of the Box CD enzyme reveals regulation of rRNA methylation
Authors	:	Lapinaite, A.; Simon, B.; Skjaerven, L.; Rakwalska-Bange, M.; Gabel, F.; Carlomagno, T.
Deposited on	:	2013-07-18

This is a Full wwPDB NMR Structure Validation 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/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

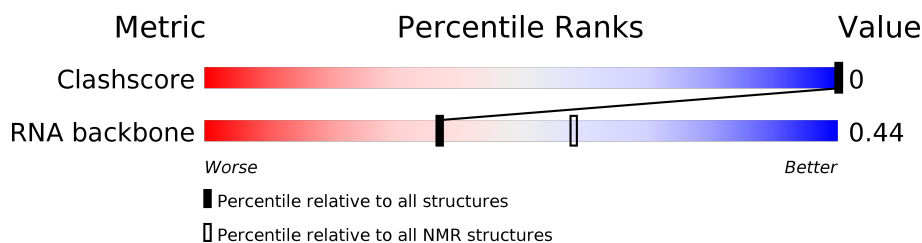|		
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 0%.

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	Percentile Ranks	Value
Clashscore		0
RNA backbone		0.44

Worse   *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
RNA backbone	4643	676

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	72	21% 56% 21% •
1	B	72	14% 60% 24% •
2	C	366	94% 6%
2	F	366	94% 5%
2	I	366	92% 8% •
2	L	366	94% 6%
3	D	121	98% •
3	G	121	97% •
3	J	121	98% ••
3	M	121	97% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	E	227	94% 6%
4	H	227	94% 5%
4	K	227	96%
4	N	227	96%
5	W	11	9% 64% 27%
5	X	11	73% 27%
5	Y	11	9% 73% 18%
5	Z	11	9% 73% 18%

2 Ensemble composition and analysis (i)

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 52476 atoms, of which 25544 are hydrogens and 0 are deuteriums.

- Molecule 1 is a RNA chain called SSR26.

Mol	Chain	Residues	Atoms						Trace
1	A	72	Total	C	H	N	O	P	0
			2342	696	782	300	493	71	
1	B	72	Total	C	H	N	O	P	0
			2342	696	782	300	493	71	

- Molecule 2 is a protein called NOP5/NOP56 RELATED PROTEIN.

Mol	Chain	Residues	Atoms						Trace
2	C	366	Total	C	H	N	O	S	0
			6006	1903	3025	518	553	7	
2	F	366	Total	C	H	N	O	S	0
			6006	1903	3025	518	553	7	
2	I	366	Total	C	H	N	O	S	0
			6006	1903	3025	518	553	7	
2	L	366	Total	C	H	N	O	S	0
			6006	1903	3025	518	553	7	

- Molecule 3 is a protein called 50S RIBOSOMAL PROTEIN L7AE.

Mol	Chain	Residues	Atoms						Trace
3	D	121	Total	C	H	N	O	S	0
			1906	591	980	153	179	3	
3	G	121	Total	C	H	N	O	S	0
			1906	591	980	153	179	3	
3	J	121	Total	C	H	N	O	S	0
			1906	591	980	153	179	3	
3	M	121	Total	C	H	N	O	S	0
			1906	591	980	153	179	3	

- Molecule 4 is a protein called FIBRILLARIN-LIKE RRNA/TRNA 2'-O-METHYLTRANS FERASE.

Mol	Chain	Residues	Atoms						Trace
4	E	227	Total	C	H	N	O	S	0
			3691	1174	1869	312	334	2	
4	H	227	Total	C	H	N	O	S	0
			3691	1174	1869	312	334	2	

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms							Trace
4	K	227	Total	C	H	N	O	S		0
			3691	1174	1869	312	334	2		
4	N	227	Total	C	H	N	O	S		0
			3691	1174	1869	312	334	2		

- Molecule 5 is a RNA chain called 5'-R(*UP*CP*GP*CP*CP*CP*AP*UP*CP*AP*CP)-3'.

Mol	Chain	Residues	Atoms						Trace
5	W	11	Total	C	H	N	O	P	0
			345	102	121	37	75	10	
5	X	11	Total	C	H	N	O	P	0
			345	102	121	37	75	10	
5	Y	11	Total	C	H	N	O	P	0
			345	102	121	37	75	10	
5	Z	11	Total	C	H	N	O	P	0
			345	102	121	37	75	10	

4 Residue-property plots (i)

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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 outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.
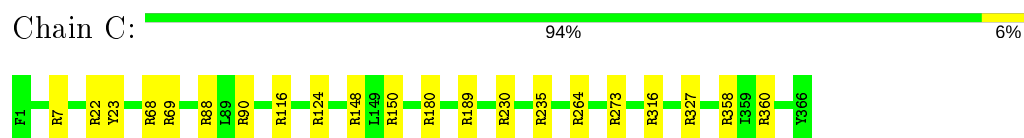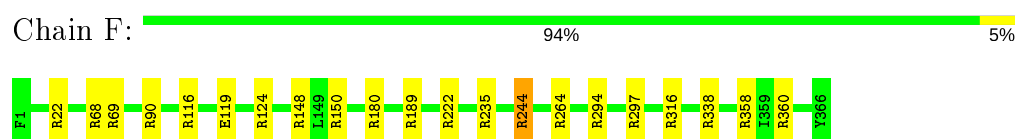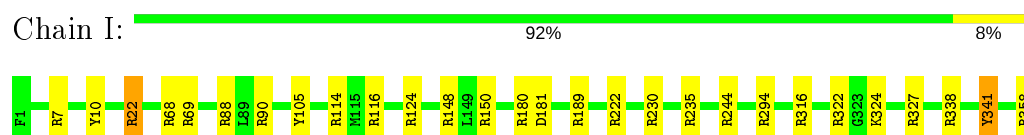
- Molecule 1: SSR26

Chain A:

- Molecule 1: SSR26

Chain B:

- Molecule 2: NOP5/NOP56 RELATED PROTEIN

Chain C:

- Molecule 2: NOP5/NOP56 RELATED PROTEIN

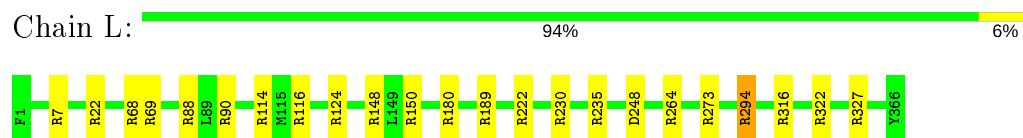Chain F:

- Molecule 2: NOP5/NOP56 RELATED PROTEIN

Chain I:

- Molecule 2: NOP5/NOP56 RELATED PROTEIN

Chain L:

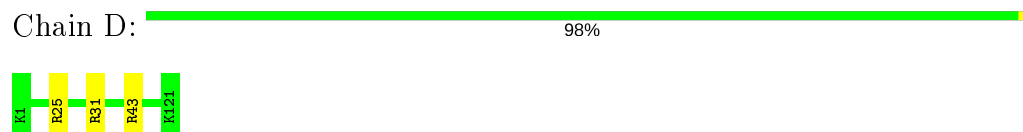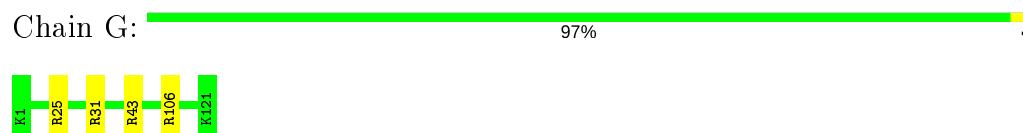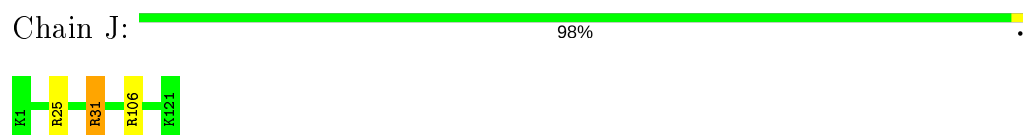94%   6%

F1 R7 R22 R68 R69 R88 L89 R90 R114 M115 R116 R124 R148 L149 R150 R180 R189 R222 R230 R235 D248 R264 R273 R294 R316 R322 R327 Y366

- Molecule 3: 50S RIBOSOMAL PROTEIN L7AE

Chain D:

98%   .

K1 R25 R31 R43 K121

- Molecule 3: 50S RIBOSOMAL PROTEIN L7AE

Chain G:

97%   .

K1 R25 R31 R43 R106 K121

- Molecule 3: 50S RIBOSOMAL PROTEIN L7AE

Chain J:

98%   ..

K1 R25 R31 R106 K121

- Molecule 3: 50S RIBOSOMAL PROTEIN L7AE

Chain M:

97%   .

K1 R25 R31 R43 R106 K121

- Molecule 4: FIBRILLARIN-LIKE RRNA/TRNA 2'-O-METHYLTRANSFERASE

Chain E:

94%   6%

M1 R34 R39 R49 R55 Y79 R109 R112 R121 R122 R138 A152 R181 R198 R225 K226 P227

- Molecule 4: FIBRILLARIN-LIKE RRNA/TRNA 2'-O-METHYLTRANSFERASE

Chain H:

94%   5% .

M1 R34 R39 R49 R55 S56 R109 R112 R121 R138 R181 R198 R209 R225 K226 P227
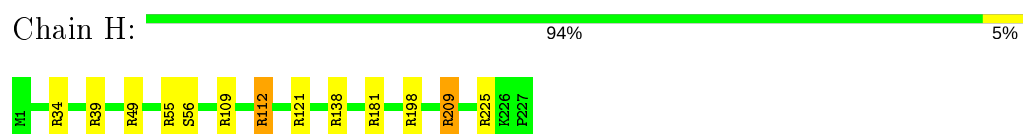
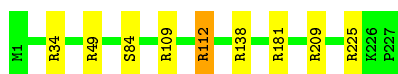- Molecule 4: FIBRILLARIN-LIKE RRNA/TRNA 2'-O-METHYLTRANSFERASE

Chain K:

96%   .

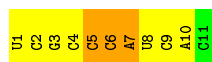• Molecule 4: FIBRILLARIN-LIKE RRNA/TRNA 2'-O-METHYLTRANSFERASE

Chain N:

• Molecule 5: 5'-R(*UP*CP*GP*CP*CP*CP*AP*UP*CP*AP*CP)-3'

Chain W:

• Molecule 5: 5'-R(*UP*CP*GP*CP*CP*CP*AP*UP*CP*AP*CP)-3'

Chain X:
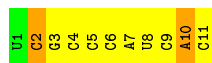

• Molecule 5: 5'-R(*UP*CP*GP*CP*CP*CP*AP*UP*CP*AP*CP)-3'

Chain Y:

• Molecule 5: 5'-R(*UP*CP*GP*CP*CP*CP*AP*UP*CP*AP*CP)-3'

Chain Z:

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *ARIA*.

Of the ? calculated structures, 1 were deposited, based on the following criterion: ?.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CNS	refinement	1.2
NMRVIEW	structure solution	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 6 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	319
Number of shifts mapped to atoms	319
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	0%

No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

5.1 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
2	I	2981	3025	3025	1
All	All	26932	25544	25544	1

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 0.

All clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
2:I:324:LYS:HE2	2:I:366:TYR:CE1	0.41	2.49

5.2 Torsion angles (i)

5.2.1 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 NMR 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	C	0	-	-	-	-
2	F	0	-	-	-	-
2	I	0	-	-	-	-
2	L	0	-	-	-	-
3	D	0	-	-	-	-
3	G	0	-	-	-	-
3	J	0	-	-	-	-
3	M	0	-	-	-	-
4	E	0	-	-	-	-
4	H	0	-	-	-	-
4	K	0	-	-	-	-
4	N	0	-	-	-	-
All	All	0	-	-	-	-

There are no Ramachandran outliers.

5.2.2 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 NMR 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	C	0	-	-	-
2	F	0	-	-	-
2	I	0	-	-	-
2	L	0	-	-	-
3	D	0	-	-	-
3	G	0	-	-	-

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	J	0	-	-	-
3	M	0	-	-	-
4	E	0	-	-	-
4	H	0	-	-	-
4	K	0	-	-	-
4	N	0	-	-	-
All	All	0	-	-	-

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

5.2.3   RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
1	A	71/72 (99%)	10 (14%)	0 (0%)	0.44
1	B	71/72 (99%)	12 (17%)	0 (0%)	0.45
5	W	10/11 (91%)	0 (0%)	0 (0%)	0.40
5	X	10/11 (91%)	0 (0%)	0 (0%)	0.46
5	Y	10/11 (91%)	0 (0%)	0 (0%)	0.48
5	Z	10/11 (91%)	1 (10%)	0 (0%)	0.35
All	All	182/188 (97%)	23 (13%)	0 (0%)	0.44

The overall RNA backbone suiteness is 0.44.

All RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	B	27	C
1	A	16	U
1	A	39	G
1	B	15	G
1	B	36	A
1	B	50	A
1	A	71	G
5	Z	10	A
1	B	16	U
1	B	39	G
1	A	62	A
1	B	63	C
1	A	15	G
1	A	63	C
1	A	26	A
1	A	14	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	27	C
1	B	14	A
1	A	35	G
1	B	37	A
1	B	26	A
1	B	38	A
1	B	71	G

There are no RNA pucker outliers to report.

5.3 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.4 Carbohydrates (i)

There are no carbohydrates in this entry.

5.5 Ligand geometry (i)

There are no ligands in this entry.

5.6 Other polymers (i)

There are no such molecules in this entry.

5.7 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Chemical shift validation ⓘ

The completeness of assignment taking into account all chemical shift lists is 0% for the well-defined parts and 0% for the entire structure.

6.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_0*

6.1.1 Bookkeeping ⓘ

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	319
Number of shifts mapped to atoms	319
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

6.1.2 Chemical shift referencing ⓘ

No chemical shift referencing corrections were calculated (not enough data).

6.1.3 Completeness of resonance assignments ⓘ

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 0%, i.e. 172 atoms were assigned a chemical shift out of a possible 40078. 0 out of 496 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	0/14008 (0%)	0/5576 (0%)	0/5712 (0%)	0/2720 (0%)
Sidechain	172/20192 (1%)	80/11816 (1%)	92/7420 (1%)	0/956 (0%)
Aromatic	0/2312 (0%)	0/1220 (0%)	0/1016 (0%)	0/76 (0%)
Overall	172/40078 (0%)	80/20674 (0%)	92/15406 (1%)	0/3998 (0%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 0%, i.e. 172 atoms were assigned a chemical shift out of a possible 40078. 0 out of 496 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	0/14008 (0%)	0/5576 (0%)	0/5712 (0%)	0/2720 (0%)
Sidechain	172/20192 (1%)	80/11816 (1%)	92/7420 (1%)	0/956 (0%)
Aromatic	0/2312 (0%)	0/1220 (0%)	0/1016 (0%)	0/76 (0%)
Overall	172/40078 (0%)	80/20674 (0%)	92/15406 (1%)	0/3998 (0%)

6.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

6.1.5 Random Coil Index (RCI) plots (i)

No *random coil index* (RCI) plot could be generated from the current chemical shift list (assigned_chem_shift_list_0). RCI is only applicable to proteins.