wwPDB X-ray Structure Validation Summary Report (i)

Aug 22, 2023 – 10:41 AM EDT

PDB ID	:	2RFK
Title	:	Substrate RNA Positioning in the Archaeal H/ACA Ribonucleoprotein Complex
Authors	:	Liang, B.; Xue, S.; Terns, R.M.; Terns, M.P.; Li, H.; Southeast Collaboratory for Structural Genomics (SECSG)
Deposited on	:	2007-09-30
Resolution	:	2.87 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

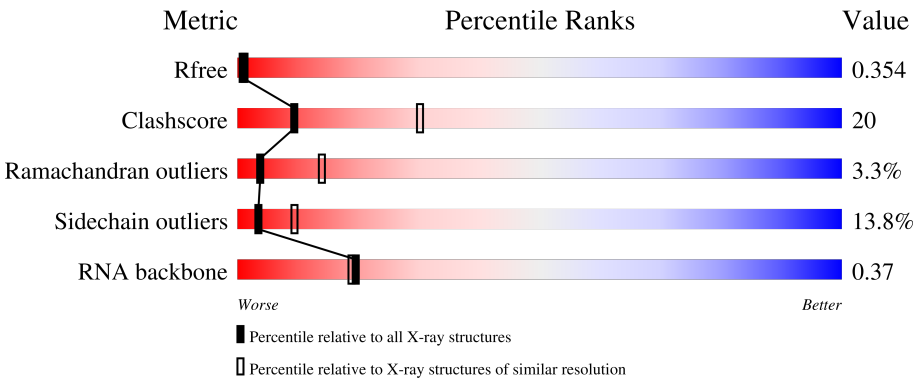
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.87 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2691 (2.90-2.86)
Clashscore	141614	2947 (2.90-2.86)
Ramachandran outliers	138981	2868 (2.90-2.86)
Sidechain outliers	138945	2871 (2.90-2.86)
RNA backbone	3102	1121 (3.16-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower 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 <=5%

Mol	Chain	Length	Quality of chain
1	D	21	43% 33% 24%
2	E	26	42% 27% 19% 12%
3	F	14	21% 36% 43%
4	A	334	54% 37% 8% •
5	B	53	58% 34% 8%
6	C	74	45% 45% 8% •

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 5005 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 guide RNA 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	D	21	450	201	86	143	20	0	0	0

- Molecule 2 is a RNA chain called guide RNA 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	E	26	543	244	94	180	25	0	0	0

- Molecule 3 is a RNA chain called target RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	F	14	300	134	54	99	13	0	0	0

- Molecule 4 is a protein called Probable tRNA pseudouridine synthase B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	A	334	2665	1716	466	473	10	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	85	ALA	ASP	engineered mutation	UNP Q7LWY0

- Molecule 5 is a protein called Ribosome biogenesis protein Nop10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	B	53	445	283	86	72	4	0	0	0

- Molecule 6 is a protein called Small nucleolar rnp similar to gar1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	C	74	601	396	102	101	2	0	0	0

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

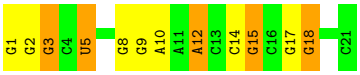
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
7	B	1	1	1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

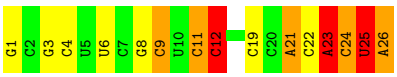
• Molecule 1: guide RNA 1

Chain D: 43% 33% 24%



• Molecule 2: guide RNA 2

Chain E: 42% 27% 19% 12%

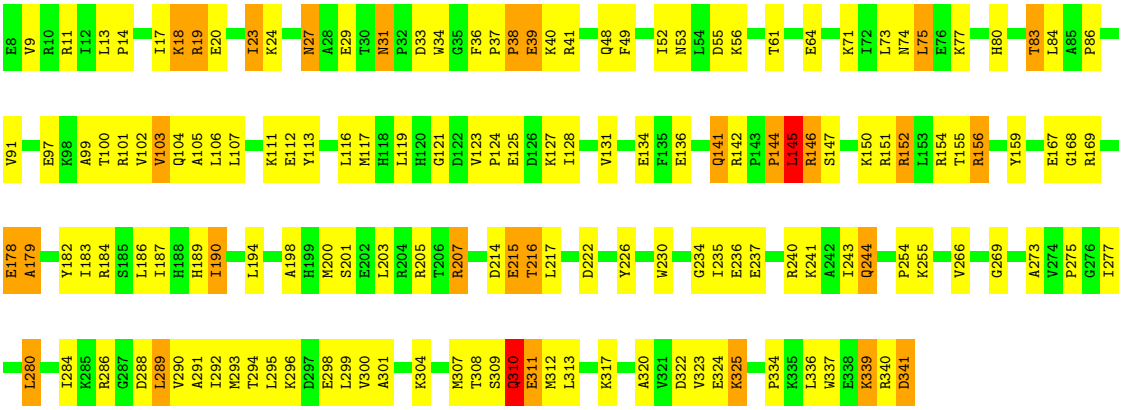


• Molecule 3: target RNA

Chain F: 21% 36% 43%

• Molecule 4: Probable tRNA pseudouridine synthase B

Chain A: 54% 37% 8%

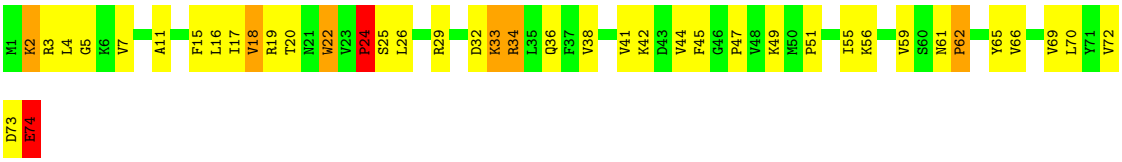


• Molecule 5: Ribosome biogenesis protein Nop10

Chain B: 58% 34% 8%

• Molecule 6: Small nucleolar rnp similar to gar1

Chain C: 45% 45% 8% •



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	96.56Å 96.56Å 240.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.52 – 2.87 38.04 – 2.79	Depositor EDS
% Data completeness (in resolution range)	80.0 (42.52-2.87) 75.5 (38.04-2.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	4.10 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.248 , 0.300 0.315 , 0.354	Depositor DCC
R_{free} test set	1175 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å^2)	67.0	Xtriage
Anisotropy	0.717	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 9.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	5005	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.43% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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

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	D	0.72	0/503	1.39	5/784 (0.6%)
2	E	0.80	0/604	1.51	7/938 (0.7%)
3	F	0.90	0/335	1.51	5/522 (1.0%)
4	A	0.55	1/2725 (0.0%)	0.67	0/3684
5	B	0.53	0/458	0.77	1/613 (0.2%)
6	C	1.33	3/615 (0.5%)	0.71	1/831 (0.1%)
All	All	0.75	4/5240 (0.1%)	0.99	19/7372 (0.3%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	74	GLU	CD-OE1	21.63	1.49	1.25
6	C	74	GLU	CD-OE2	20.98	1.48	1.25
6	C	36	GLN	CD-OE1	5.33	1.35	1.24
4	A	145	LEU	C-N	5.17	1.46	1.34

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	C	74	GLU	OE1-CD-OE2	8.35	133.32	123.30
1	D	5	U	C3'-C2'-C1'	-7.72	95.32	101.50
3	F	4	G	P-O3'-C3'	6.72	127.76	119.70
5	B	23	CYS	CA-CB-SG	6.27	125.29	114.00
2	E	12	C	C3'-C2'-C1'	-6.09	96.63	101.50

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	D	450	0	232	8	0
2	E	543	0	284	14	0
3	F	300	0	152	9	0
4	A	2665	0	2750	116	0
5	B	445	0	451	19	0
6	C	601	0	637	34	0
7	B	1	0	0	0	0
All	All	5005	0	4506	186	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 20.

The worst 5 of 186 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)
5:B:20:CYS:HB3	5:B:23:CYS:HB3	1.32	1.08
4:A:284:ILE:HB	4:A:309:SER:HB3	1.46	0.95
3:F:10:U:H4'	3:F:11:G:O5'	1.66	0.92
4:A:19:ARG:HH12	4:A:254:PRO:HB3	1.36	0.88
4:A:41:ARG:O	4:A:240:ARG:NH2	2.09	0.86

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 X-ray entries followed by that with respect to entries of similar resolution.

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
4	A	332/334 (99%)	281 (85%)	40 (12%)	11 (3%)	4 14
5	B	51/53 (96%)	43 (84%)	7 (14%)	1 (2%)	7 25
6	C	72/74 (97%)	60 (83%)	9 (12%)	3 (4%)	3 9
All	All	455/461 (99%)	384 (84%)	56 (12%)	15 (3%)	4 14

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	38	PRO
4	A	103	VAL
4	A	152	ARG
4	A	179	ALA
4	A	39	GLU

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 X-ray entries followed by that with respect to entries of similar resolution.

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
4	A	283/283 (100%)	244 (86%)	39 (14%)	3 9
5	B	48/48 (100%)	44 (92%)	4 (8%)	11 30
6	C	67/67 (100%)	55 (82%)	12 (18%)	2 4
All	All	398/398 (100%)	343 (86%)	55 (14%)	3 9

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	241	LYS
4	A	311	GLU
6	C	74	GLU
6	C	26	LEU
4	A	244	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
4	A	80	HIS
4	A	268	HIS
6	C	36	GLN
6	C	27	ASN
4	A	63	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	D	20/21 (95%)	7 (35%)	0
2	E	25/26 (96%)	9 (36%)	1 (4%)
3	F	14/14 (100%)	5 (35%)	2 (14%)
All	All	59/61 (96%)	21 (35%)	3 (5%)

5 of 21 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	D	3	G
1	D	5	U
1	D	8	G
1	D	10	A
1	D	12	A

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	E	23	A
3	F	4	G
3	F	10	U

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)

Of 1 ligands modelled in this entry, 1 is 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 (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.