wwPDB EM Validation Summary Report (i)

May 5, 2024 – 12:06 AM JST

PDB ID	:	8IAZ
EMDB ID	:	EMD-35323
Title	:	Cryo-EM structure of the ISFba1 TnpB-reRNA-dsDNA complex
Authors	:	Yin, M.; Zhou, F.; Zhu, Y.; Huang, Z.
Deposited on	:	2023-02-09
Resolution	:	3.00 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 (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 (i)) were used in the production of this report:

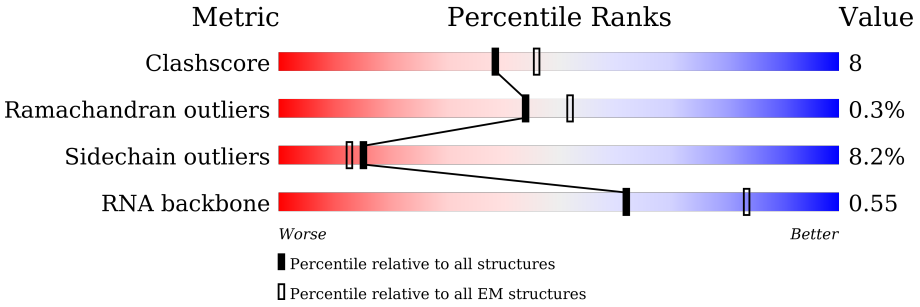
EMDB validation analysis	:	FAILED
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.00 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)
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 <=5%

Mol	Chain	Length	Quality of chain
1	A	384	78% 19% •
2	E	207	43% 28% 7% 21%
3	B	24	21% 79%
4	C	16	31% 69%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7396 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 Transposase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	384	3110	1974	565	555	16	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	371	ALA	ASP	conflict	UNP A0A417B524

- Molecule 2 is a RNA chain called RNA (207-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	E	163	3464	1555	615	1131	163	0	0

- Molecule 3 is a DNA chain called DNA (5'-D(P*AP*CP*AP*TP*GP*GP*AP*CP*CP*AP*TP*CP*AP*GP*CP*TP*CP*CP*TP*AP*AP*TP*GP*G)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	B	24	490	233	91	142	24	0	0

- Molecule 4 is a DNA chain called DNA (5'-D(P*CP*CP*AP*TP*TP*AP*GP*GP*AP*GP*CP*TP*GP*AP*TP*G)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
4	C	16	331	157	62	96	16	0	0

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
5	A	1	1	1	0

3 Residue-property plots (i)

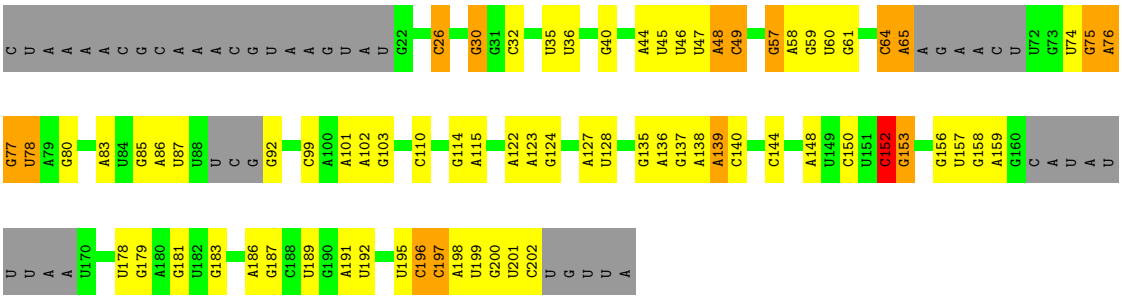
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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: Transposase

Chain A: 78% 19% •

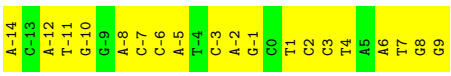
• Molecule 2: RNA (207-MER)

Chain E: 43% 28% 7% 21%



• Molecule 3: DNA (5'-D(P*AP*CP*AP*TP*GP*GP*AP*CP*CP*AP*TP*CP*AP*GP*CP*TP*CP*CP*TP*AP*AP*TP*GP*G)-3')

Chain B: 21% 79%



• Molecule 4: DNA (5'-D(P*CP*CP*AP*TP*TP*AP*GP*GP*AP*GP*CP*TP*GP*AP*TP*G)-3')

Chain C: 31% 69%

C-5
C-4
A-3
A0
G1
G2
A3
G4
C5
T6
G7
G10

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	369379	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 ($e^-/Å^2$)	40	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	1500	Depositor
Magnification	Not provided	
Image detector	FEI FALCON III (4k 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

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	0/3168	0.51	1/4243 (0.0%)
2	E	0.59	4/3874 (0.1%)	0.91	9/6016 (0.1%)
3	B	0.63	0/549	0.85	0/844
4	C	1.03	1/371 (0.3%)	1.07	3/571 (0.5%)
All	All	0.53	5/7962 (0.1%)	0.79	13/11674 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	196	C	O3'-P	-14.59	1.43	1.61
2	E	196	C	P-OP2	-7.16	1.36	1.49
2	E	197	C	P-OP2	-5.47	1.39	1.49
4	C	3	DA	O3'-P	-5.30	1.54	1.61
2	E	196	C	P-OP1	-5.10	1.40	1.49

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	197	C	O5'-P-OP1	-7.58	98.88	105.70
2	E	197	C	C1'-C2'-O2'	-7.17	89.08	110.60
4	C	4	DG	C1'-O4'-C4'	-6.83	103.27	110.10
1	A	168	ASP	CB-CG-OD1	6.30	123.97	118.30
2	E	152	C	OP1-P-O3'	5.80	117.95	105.20

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	3110	0	3197	51	0
2	E	3464	0	1747	35	0
3	B	490	0	270	18	0
4	C	331	0	181	11	0
5	A	1	0	0	0	0
All	All	7396	0	5395	97	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 8.

The worst 5 of 97 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:E:44:A:H8	2:E:110:C:HO2'	1.20	0.84
2:E:26:C:H5	2:E:40:G:H1	1.31	0.78
1:A:202:ILE:HG21	1:A:371:ALA:HB1	1.68	0.75
1:A:130:ARG:NH2	2:E:181:G:N7	2.43	0.66
1:A:80:LYS:NZ	3:B:1:DT:O4	2.29	0.64

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
1	A	382/384 (100%)	354 (93%)	27 (7%)	1 (0%)	41 76

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	132	ARG

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	A	343/347 (99%)	315 (92%)	28 (8%)	11 39

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	164	SER
1	A	368	GLU
1	A	182	LYS
1	A	324	ASN
1	A	175	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	23	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	E	159/207 (76%)	36 (22%)	7 (4%)

5 of 36 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	E	30	G
2	E	35	U
2	E	36	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	E	45	U
2	E	46	U

5 of 7 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	E	77	G
2	E	122	A
2	E	152	C
2	E	139	A
2	E	64	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)

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.