wwPDB EM Validation Summary Report (i)

Mar 31, 2024 – 12:48 AM JST

PDB ID	:	8J8H
EMDB ID	:	EMD-36070
Title	:	SPARTA monomer bound with guide-target, state 2
Authors	:	Li, Z.X.; Guo, L.J.; Huang, P.P.; Xiao, Y.B.; Chen, M.R.
Deposited on	:	2023-05-01
Resolution	:	3.40 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:

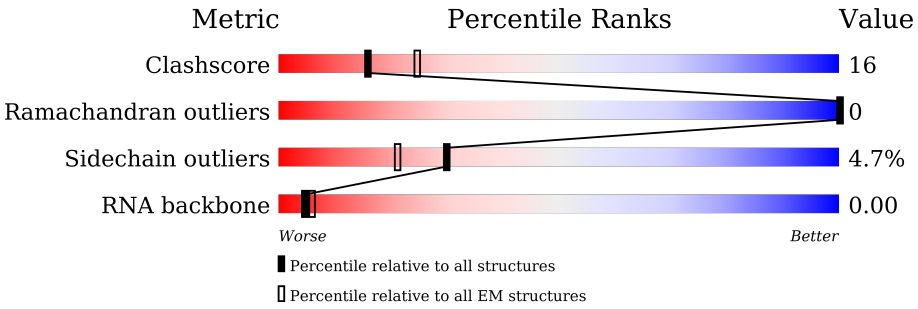
EMDB validation analysis	:	FAILED
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.40 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	541	53% 29% • 16%
2	B	484	34% 23% • 41%
3	E	21	38% 52% 5% 5%
4	F	25	24% 52% 24%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6868 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 Piwi domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	452	3651	2367	609	664	11	0	0

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-33	MET	-	initiating methionine	UNP A0A1I7NFD7
A	-32	GLY	-	expression tag	UNP A0A1I7NFD7
A	-31	SER	-	expression tag	UNP A0A1I7NFD7
A	-30	SER	-	expression tag	UNP A0A1I7NFD7
A	-29	HIS	-	expression tag	UNP A0A1I7NFD7
A	-28	HIS	-	expression tag	UNP A0A1I7NFD7
A	-27	HIS	-	expression tag	UNP A0A1I7NFD7
A	-26	HIS	-	expression tag	UNP A0A1I7NFD7
A	-25	HIS	-	expression tag	UNP A0A1I7NFD7
A	-24	HIS	-	expression tag	UNP A0A1I7NFD7
A	-23	SER	-	expression tag	UNP A0A1I7NFD7
A	-22	SER	-	expression tag	UNP A0A1I7NFD7
A	-21	GLY	-	expression tag	UNP A0A1I7NFD7
A	-20	LEU	-	expression tag	UNP A0A1I7NFD7
A	-19	VAL	-	expression tag	UNP A0A1I7NFD7
A	-18	PRO	-	expression tag	UNP A0A1I7NFD7
A	-17	ARG	-	expression tag	UNP A0A1I7NFD7
A	-16	GLY	-	expression tag	UNP A0A1I7NFD7
A	-15	SER	-	expression tag	UNP A0A1I7NFD7
A	-14	HIS	-	expression tag	UNP A0A1I7NFD7
A	-13	MET	-	expression tag	UNP A0A1I7NFD7
A	-12	ALA	-	expression tag	UNP A0A1I7NFD7
A	-11	SER	-	expression tag	UNP A0A1I7NFD7
A	-10	MET	-	expression tag	UNP A0A1I7NFD7
A	-9	THR	-	expression tag	UNP A0A1I7NFD7
A	-8	GLY	-	expression tag	UNP A0A1I7NFD7
A	-7	GLY	-	expression tag	UNP A0A1I7NFD7
A	-6	GLN	-	expression tag	UNP A0A1I7NFD7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLN	-	expression tag	UNP A0A1I7NFD7
A	-4	MET	-	expression tag	UNP A0A1I7NFD7
A	-3	GLY	-	expression tag	UNP A0A1I7NFD7
A	-2	ARG	-	expression tag	UNP A0A1I7NFD7
A	-1	GLY	-	expression tag	UNP A0A1I7NFD7
A	0	SER	-	expression tag	UNP A0A1I7NFD7

- Molecule 2 is a protein called TIR domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	286	Total	C	N	O	S	0	0
			2421	1573	413	428	7		

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-33	MET	-	initiating methionine	UNP A0A1I7NFG5
B	-32	GLY	-	expression tag	UNP A0A1I7NFG5
B	-31	SER	-	expression tag	UNP A0A1I7NFG5
B	-30	SER	-	expression tag	UNP A0A1I7NFG5
B	-29	HIS	-	expression tag	UNP A0A1I7NFG5
B	-28	HIS	-	expression tag	UNP A0A1I7NFG5
B	-27	HIS	-	expression tag	UNP A0A1I7NFG5
B	-26	HIS	-	expression tag	UNP A0A1I7NFG5
B	-25	HIS	-	expression tag	UNP A0A1I7NFG5
B	-24	HIS	-	expression tag	UNP A0A1I7NFG5
B	-23	SER	-	expression tag	UNP A0A1I7NFG5
B	-22	SER	-	expression tag	UNP A0A1I7NFG5
B	-21	GLY	-	expression tag	UNP A0A1I7NFG5
B	-20	LEU	-	expression tag	UNP A0A1I7NFG5
B	-19	VAL	-	expression tag	UNP A0A1I7NFG5
B	-18	PRO	-	expression tag	UNP A0A1I7NFG5
B	-17	ARG	-	expression tag	UNP A0A1I7NFG5
B	-16	GLY	-	expression tag	UNP A0A1I7NFG5
B	-15	SER	-	expression tag	UNP A0A1I7NFG5
B	-14	HIS	-	expression tag	UNP A0A1I7NFG5
B	-13	MET	-	expression tag	UNP A0A1I7NFG5
B	-12	ALA	-	expression tag	UNP A0A1I7NFG5
B	-11	SER	-	expression tag	UNP A0A1I7NFG5
B	-10	MET	-	expression tag	UNP A0A1I7NFG5
B	-9	THR	-	expression tag	UNP A0A1I7NFG5
B	-8	GLY	-	expression tag	UNP A0A1I7NFG5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-7	GLY	-	expression tag	UNP A0A1I7NFG5
B	-6	GLN	-	expression tag	UNP A0A1I7NFG5
B	-5	GLN	-	expression tag	UNP A0A1I7NFG5
B	-4	MET	-	expression tag	UNP A0A1I7NFG5
B	-3	GLY	-	expression tag	UNP A0A1I7NFG5
B	-2	ARG	-	expression tag	UNP A0A1I7NFG5
B	-1	GLY	-	expression tag	UNP A0A1I7NFG5
B	0	SER	-	expression tag	UNP A0A1I7NFG5

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	E	20	Total	C	N	O	P	0	0
			403	189	71	123	20		

- Molecule 4 is a DNA chain called DNA (25-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
4	F	19	Total	C	N	O	P	3	0
			392	187	74	112	19		

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

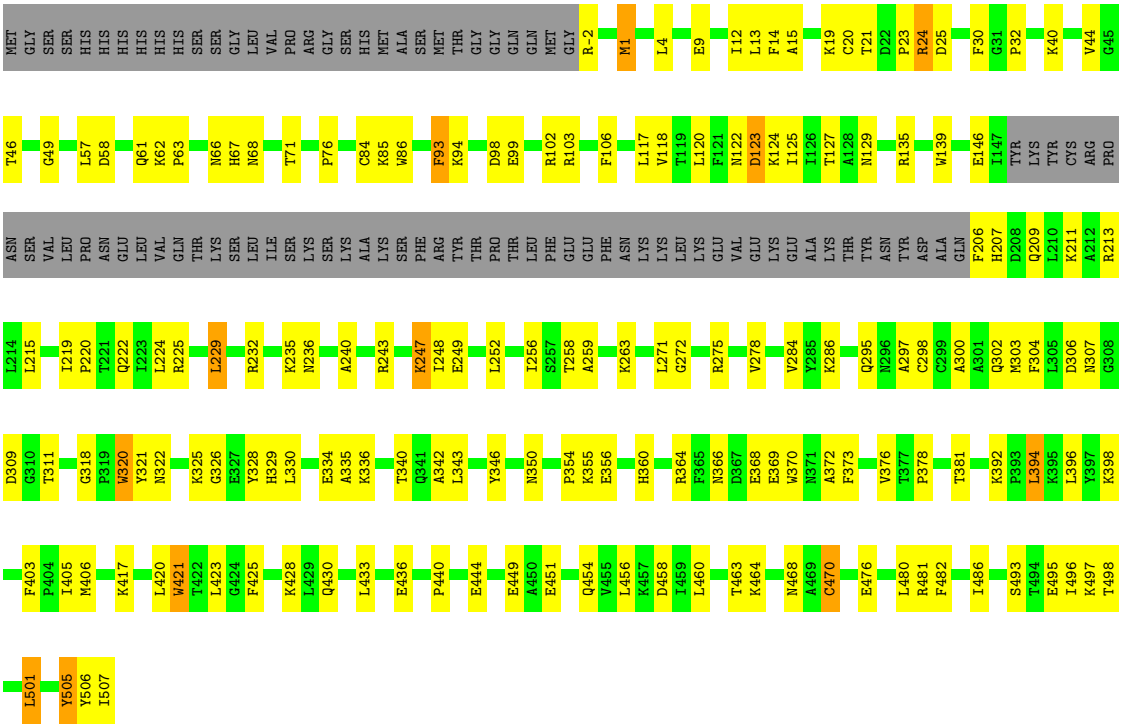
Mol	Chain	Residues	Atoms		AltConf
5	A	1	Total	Mg	0
			1	1	

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.

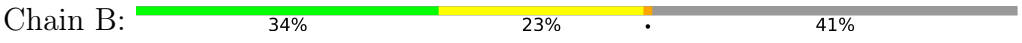
• Molecule 1: Piwi domain-containing protein

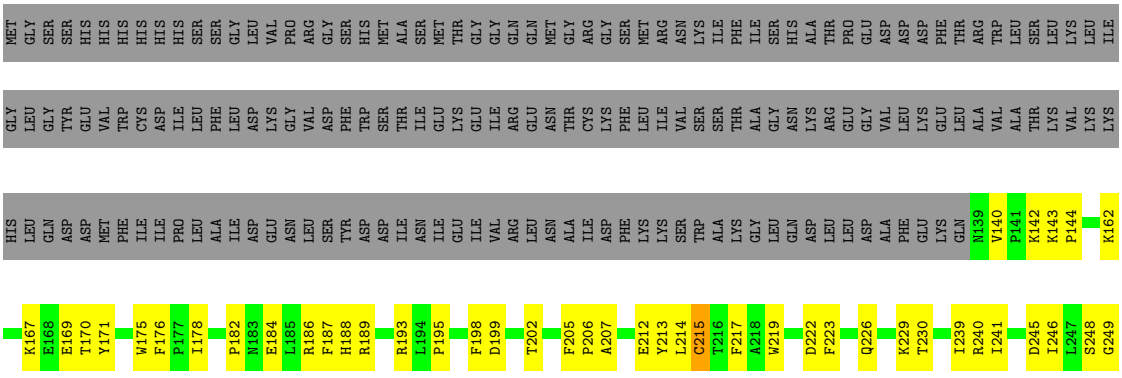
Chain A:

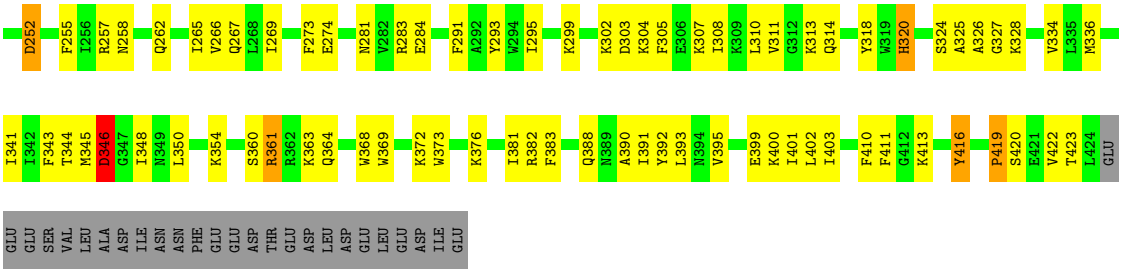


• Molecule 2: TIR domain-containing protein

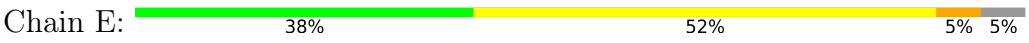
Chain B:

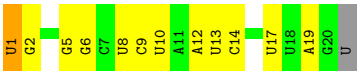




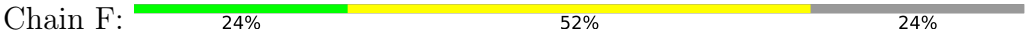


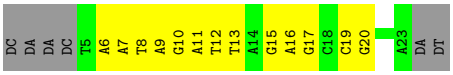
• Molecule 3: RNA (5'-R(P*UP*GP*AP*CP*GP*GP*CP*UP*CP*UP*AP*AP*UP*CP*UP*AP*UP*UP*AP*GP*U)-3')





• Molecule 4: DNA (25-MER)





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	117819	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)	1000	Depositor
Maximum defocus (nm)	2400	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG

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	0.27	0/3743	0.53	2/5066 (0.0%)
2	B	0.27	0/2488	0.54	2/3352 (0.1%)
3	E	0.62	1/451 (0.2%)	0.78	2/689 (0.3%)
4	F	0.52	0/440	0.88	0/677
All	All	0.33	1/7122 (0.0%)	0.58	6/9784 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	1	U	OP3-P	-10.67	1.48	1.61

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	419	PRO	CA-N-CD	-8.29	99.90	111.50
3	E	17	U	O4'-C1'-N1	7.30	114.04	108.20
1	A	394	LEU	CA-CB-CG	6.76	130.85	115.30
3	E	1	U	O4'-C1'-N1	6.75	113.60	108.20
1	A	215	LEU	CA-CB-CG	6.00	129.09	115.30

There are no chirality outliers.

There are no planarity outliers.

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	A	3651	0	3673	120	0
2	B	2421	0	2422	83	0
3	E	403	0	214	14	0
4	F	392	0	201	13	0
5	A	1	0	0	0	0
All	All	6868	0	6510	210	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 16.

The worst 5 of 210 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:B:195:PRO:HG2	2:B:198:PHE:HB2	1.61	0.80
2:B:364:GLN:HE21	2:B:368:TRP:HE1	1.32	0.76
1:A:311:THR:HG23	1:A:501:LEU:HB2	1.72	0.71
1:A:328:TYR:O	1:A:329:HIS:ND1	2.25	0.69
4:F:12[A]:DT:H2"	4:F:13[A]:DT:H5"	1.73	0.69

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	448/541 (83%)	417 (93%)	31 (7%)	0	100	100
2	B	284/484 (59%)	253 (89%)	31 (11%)	0	100	100
All	All	732/1025 (71%)	670 (92%)	62 (8%)	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
1	A	393/472 (83%)	377 (96%)	16 (4%)	30 59
2	B	266/441 (60%)	251 (94%)	15 (6%)	21 51
All	All	659/913 (72%)	628 (95%)	31 (5%)	30 57

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	501	LEU
2	B	346	ASP
2	B	215	CYS
2	B	369	TRP
2	B	291	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	322	ASN
1	A	454	GLN
2	B	364	GLN
1	A	68	ASN
1	A	66	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	E	0/21	-	-

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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.