



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 21, 2021 – 04:11 PM JST

PDB ID : 7EL1
Title : Structure of a protein from bacteria
Authors : Liu, H.; Zhu, Y.; Huang, Z.
Deposited on : 2021-04-07
Resolution : 2.22 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.22
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)	:	2.22

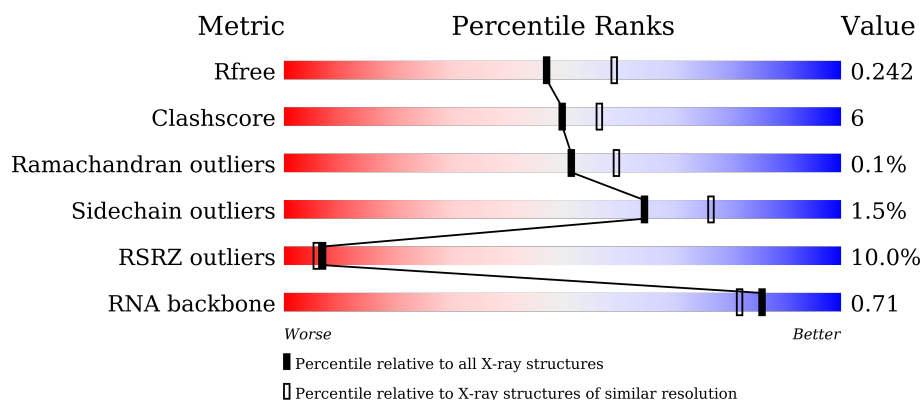
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.22 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	5912 (2.24-2.20)
Clashscore	141614	6646 (2.24-2.20)
Ramachandran outliers	138981	6543 (2.24-2.20)
Sidechain outliers	138945	6544 (2.24-2.20)
RSRZ outliers	127900	5797 (2.24-2.20)
RNA backbone	3102	1049 (2.64-1.80)

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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1053	<div> <div>11%</div> <div>85%</div> <div>13%</div> <div>.</div> </div>
2	B	73	<div> <div>11%</div> <div>78%</div> <div>18%</div> <div>..</div> </div>
3	C	28	<div> <div>89%</div> <div>7%</div> <div>.</div> </div>
4	D	8	<div> <div>88%</div> <div>12%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	100	<div><div></div><div>4%</div><div>79%</div><div>17%</div><div>••</div></div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11628 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 protein called CRISPR-associated endonuclease Cas9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1033	Total	C	N	O	S	0	0	0
			8413	5349	1466	1586	12			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	580	ALA	ASN	engineered mutation	UNP J7RUA5
A	946	ALA	CYS	engineered mutation	UNP J7RUA5

- Molecule 2 is a RNA chain called RNA (73-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	73	Total	C	N	O	P	0	0	0
			1545	689	279	504	73			

- Molecule 3 is a DNA chain called DNA (28-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	28	Total	C	N	O	P	0	0	0
			561	270	99	165	27			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	8	Total	C	N	O	P	0	0	0
			164	80	31	46	7			

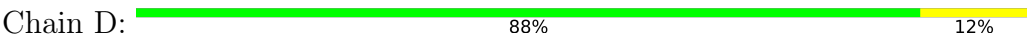
- Molecule 5 is a protein called 100AA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	98	Total 819	C 526	N 134	O 157	S 2	0	0	0

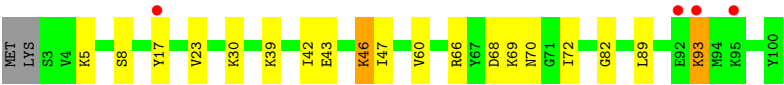
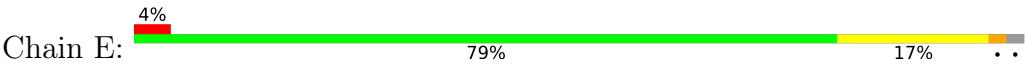
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	73	Total 73	O 73	0	0
6	B	35	Total 35	O 35	0	0
6	C	5	Total 5	O 5	0	0
6	D	9	Total 9	O 9	0	0
6	E	4	Total 4	O 4	0	0

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



● Molecule 5: 100AA



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	330.58Å 105.22Å 68.55Å 90.00° 92.71° 90.00°	Depositor
Resolution (Å)	50.00 – 2.22 44.37 – 2.22	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.00-2.22) 99.8 (44.37-2.22)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.43 (at 2.22Å)	Xtriage
Refinement program	PHENIX v1.11	Depositor
R, R_{free}	0.213 , 0.242 0.213 , 0.242	Depositor DCC
R_{free} test set	5675 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å ²)	54.9	Xtriage
Anisotropy	0.166	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 50.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.011 for -h,-k,l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	11628	wwPDB-VP
Average B, all atoms (Å ²)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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 > 5$	RMSZ	# $ Z > 5$
1	A	0.47	1/8559 (0.0%)	0.64	5/11529 (0.0%)
2	B	0.71	1/1729 (0.1%)	1.15	3/2692 (0.1%)
3	C	1.02	0/627	1.18	2/963 (0.2%)
4	D	1.26	0/184	1.35	2/283 (0.7%)
5	E	0.51	0/836	0.73	1/1120 (0.1%)
All	All	0.57	2/11935 (0.0%)	0.80	13/16587 (0.1%)

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	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	G	OP3-P	-9.19	1.50	1.61
1	A	308	GLU	CB-CG	5.27	1.62	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	57	G	O5'-P-OP2	-14.63	92.53	105.70
5	E	46	LYS	CD-CE-NZ	-12.60	82.72	111.70
1	A	878	LYS	CD-CE-NZ	-10.50	87.55	111.70
4	D	4	DA	O4'-C1'-N9	8.57	114.00	108.00
1	A	208	ARG	NE-CZ-NH2	7.01	123.80	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	480	ARG	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	A	8413	0	8422	113	0
2	B	1545	0	773	6	0
3	C	561	0	317	2	0
4	D	164	0	93	0	0
5	E	819	0	801	11	0
6	A	73	0	0	2	0
6	B	35	0	0	0	0
6	C	5	0	0	0	0
6	D	9	0	0	0	0
6	E	4	0	0	1	0
All	All	11628	0	10406	126	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 6.

The worst 5 of 126 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)
1:A:520:LEU:O	1:A:524:ILE:HD12	1.55	1.06
1:A:134:THR:HG22	1:A:170:ARG:HH11	1.28	0.96
1:A:520:LEU:HD13	1:A:524:ILE:HD11	1.48	0.95
1:A:521:ILE:HA	1:A:524:ILE:CD1	1.97	0.94
1:A:504:ARG:HA	1:A:504:ARG:HH21	1.36	0.88

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 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	
1	A	1029/1053 (98%)	992 (96%)	36 (4%)	1 (0%)	51	60
5	E	96/100 (96%)	94 (98%)	2 (2%)	0	100	100
All	All	1125/1153 (98%)	1086 (96%)	38 (3%)	1 (0%)	51	60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	482	LYS

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 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	
1	A	911/970 (94%)	898 (99%)	13 (1%)	67	78
5	E	91/94 (97%)	89 (98%)	2 (2%)	52	64
All	All	1002/1064 (94%)	987 (98%)	15 (2%)	65	76

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	387	LYS
5	E	46	LYS
1	A	419	ASN
5	E	93	LYS
1	A	617	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	570	ASN
1	A	700	HIS
1	A	985	ASN
1	A	835	GLN
1	A	433	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	71/73 (97%)	6 (8%)	1 (1%)

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	38	A
2	B	39	C
2	B	55	A
2	B	57	G
2	B	63	A

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	B	62	A

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](#)

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](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	1033/1053 (98%)	0.80	112 (10%) 5 4	36, 65, 103, 146	0
2	B	73/73 (100%)	0.85	8 (10%) 5 4	42, 54, 112, 120	0
3	C	28/28 (100%)	0.12	0 100 100	46, 56, 82, 83	0
4	D	8/8 (100%)	-0.08	0 100 100	41, 45, 50, 52	0
5	E	98/100 (98%)	0.67	4 (4%) 37 35	53, 69, 81, 82	0
All	All	1240/1262 (98%)	0.77	124 (10%) 7 6	36, 65, 101, 146	0

The worst 5 of 124 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	614	GLY	16.6
1	A	492	ASN	11.1
1	A	615	LYS	9.8
1	A	505	ILE	8.9
1	A	494	MET	8.9

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

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.