



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 16, 2023 – 10:54 AM EST

PDB ID : 4P0R
Title : human Mus81-Eme1-3'flap DNA complex
Authors : Gwon, G.H.; Baek, K.; Cho, Y.
Deposited on : 2014-02-22
Resolution : 6.50 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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 : 2.36
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) : 2.36

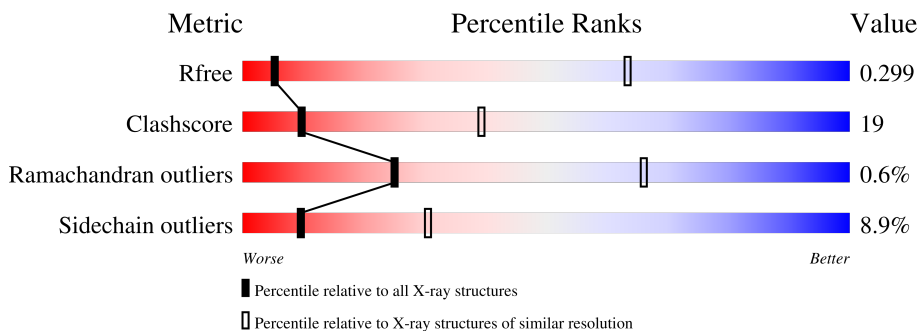
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 6.50 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	1000 (9.00-3.90)
Clashscore	141614	1064 (9.00-3.90)
Ramachandran outliers	138981	1012 (9.00-3.88)
Sidechain outliers	138945	1010 (9.00-3.84)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	306	51% (green), 36% (yellow), 10% (grey)
1	C	306	52% (green), 33% (yellow), 10% (grey)
2	B	393	48% (green), 21% (yellow), 28% (grey)
2	D	393	47% (green), 22% (yellow), 28% (grey)
3	E	15	33% (green), 47% (yellow), 20% (grey)
3	H	15	20% (green), 60% (yellow), 20% (grey)
4	F	32	9% (green), 53% (yellow), 9% (orange), 28% (grey)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	I	32	
5	G	20	
5	J	20	

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 10646 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 Crossover junction endonuclease MUS81.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	275	Total	C	N	O	S	0	0	0
			2173	1364	403	398	8			
1	C	275	Total	C	N	O	S	0	0	0
			2173	1364	403	398	8			

- Molecule 2 is a protein called Crossover junction endonuclease EME1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	284	Total	C	N	O	S	0	0	0
			2227	1401	397	415	14			
2	D	284	Total	C	N	O	S	0	0	0
			2227	1401	397	415	14			

- Molecule 3 is a DNA chain called DNA CTGTGTGTAAGCACG.

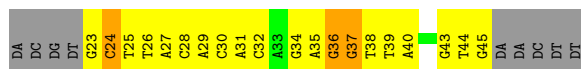
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	E	12	Total	C	N	O	P	0	0	0
			248	118	44	74	12			
3	H	12	Total	C	N	O	P	0	0	0
			248	118	44	74	12			

- Molecule 4 is a DNA chain called DNA ACGTGCTTACACACAGAGGTTAGGGTGAAC TT.

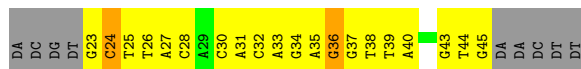
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	F	23	Total	C	N	O	P	0	0	0
			478	226	92	137	23			
4	I	23	Total	C	N	O	P	0	0	0
			478	226	92	137	23			

- Molecule 5 is a DNA chain called DNA CAAGTTCACCCTAACCTCAG.

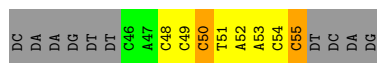
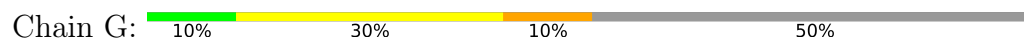
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
5	G	10	197	94	35	58	10	0	0	0
5	J	10	197	94	35	58	10	0	0	0



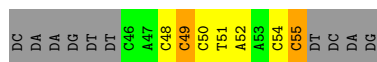
- Molecule 4: DNA ACGTGCTTACACACAGAGGTTAGGGTGAACTT



- Molecule 5: DNA CAAGTTCACCCTAACCTCAG



- Molecule 5: DNA CAAGTTCACCCTAACCTCAG



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	221.76Å 135.30Å 102.89Å 90.00° 113.27° 90.00°	Depositor
Resolution (Å)	34.86 – 6.50 47.44 – 5.97	Depositor EDS
% Data completeness (in resolution range)	99.1 (34.86-6.50) 99.2 (47.44-5.97)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.39 (at 6.15Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.210 , 0.295 0.214 , 0.299	Depositor DCC
R_{free} test set	335 reflections (4.68%)	wwPDB-VP
Wilson B-factor (Å ²)	371.5	Xtrriage
Anisotropy	0.339	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 93.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.84	EDS
Total number of atoms	10646	wwPDB-VP
Average B, all atoms (Å ²)	163.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.27% 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 i

5.1 Standard geometry i

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	0.58	0/2211	0.71	0/2989
1	C	0.57	0/2211	0.74	1/2989 (0.0%)
2	B	0.63	0/2258	0.92	5/3054 (0.2%)
2	D	0.66	1/2258 (0.0%)	0.93	5/3054 (0.2%)
3	E	0.63	0/277	1.12	0/426
3	H	0.69	0/277	1.11	0/426
4	F	0.75	0/537	1.38	4/828 (0.5%)
4	I	0.73	0/537	1.34	2/828 (0.2%)
5	G	0.72	0/219	1.35	2/333 (0.6%)
5	J	0.72	0/219	1.39	4/333 (1.2%)
All	All	0.63	1/11004 (0.0%)	0.95	23/15260 (0.2%)

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
2	B	0	1
2	D	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	246	GLU	CG-CD	-6.27	1.42	1.51

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	484	ARG	NE-CZ-NH2	-9.31	115.65	120.30
4	I	36	DG	O4'-C1'-N9	7.88	113.51	108.00
4	F	36	DG	O4'-C1'-N9	7.60	113.32	108.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	J	50	DC	O4'-C1'-N1	7.50	113.25	108.00
4	F	24	DC	O4'-C1'-N1	7.36	113.15	108.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	467	GLY	Peptide
2	D	544	ARG	Mainchain

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	2173	0	2203	97	0
1	C	2173	0	2203	102	0
2	B	2227	0	2265	84	0
2	D	2227	0	2265	92	0
3	E	248	0	137	11	0
3	H	248	0	137	13	0
4	F	478	0	259	34	0
4	I	478	0	259	33	0
5	G	197	0	112	15	0
5	J	197	0	112	11	0
All	All	10646	0	9952	389	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 19.

The worst 5 of 389 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:C:486:SER:H	1:C:489:LYS:HB2	1.34	0.92
2:D:490:ASN:O	2:D:552:ARG:NH2	2.05	0.90
1:C:472:ARG:HB2	2:D:562:PRO:HB3	1.55	0.88

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:490:ASN:O	2:B:552:ARG:NH2	2.06	0.87
2:B:493:SER:HB3	5:G:50:DC:H3'	1.56	0.87

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
1	A	267/306 (87%)	243 (91%)	22 (8%)	2 (1%)	22 63
1	C	267/306 (87%)	244 (91%)	21 (8%)	2 (1%)	22 63
2	B	276/393 (70%)	261 (95%)	14 (5%)	1 (0%)	34 72
2	D	276/393 (70%)	260 (94%)	14 (5%)	2 (1%)	22 63
All	All	1086/1398 (78%)	1008 (93%)	71 (6%)	7 (1%)	25 66

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	259	PRO
1	C	259	PRO
2	D	562	PRO
2	D	467	GLY
1	C	432	PRO

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	
1	A	237/259 (92%)	216 (91%)	21 (9%)	9	30
1	C	237/259 (92%)	210 (89%)	27 (11%)	5	21
2	B	242/334 (72%)	224 (93%)	18 (7%)	13	38
2	D	242/334 (72%)	223 (92%)	19 (8%)	12	36
All	All	958/1186 (81%)	873 (91%)	85 (9%)	9	30

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	474	VAL
2	D	287	VAL
1	C	482	VAL
1	C	550	LEU
2	D	342	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	262	GLN
2	D	488	GLN
2	D	556	GLN
2	B	262	GLN
1	C	322	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

6.1 Protein, DNA and RNA chains

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

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

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

6.4 Ligands

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

6.5 Other polymers

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