



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 2, 2026 – 02:43 AM UTC

PDB ID : 4EJZ / pdb_00004ejz
Title : Structure of MBOgg1 in complex with low affinity DNA ligand
Authors : Jiang, T.; Yu, H.J.; Bi, L.J.; Zhang, X.E.; Yang, M.Z.
Deposited on : 2012-04-08
Resolution : 3.05 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

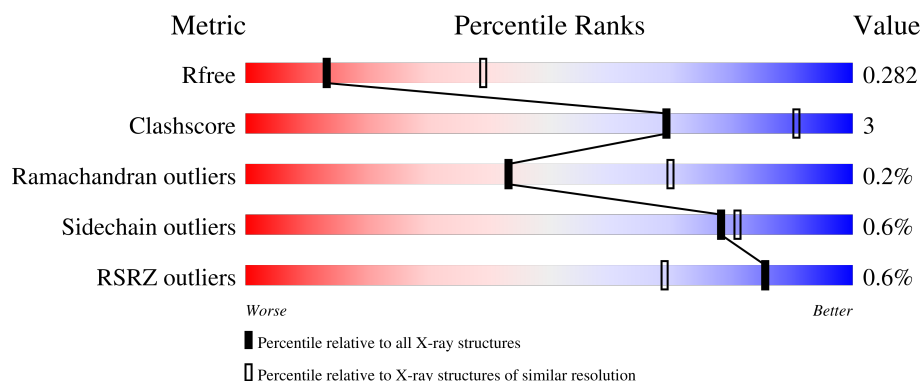
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

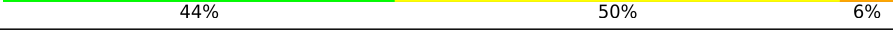
The reported resolution of this entry is 3.05 Å.

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}	180053	2469 (3.10-3.02)
Clashscore	190562	2569 (3.10-3.02)
Ramachandran outliers	187476	2424 (3.10-3.02)
Sidechain outliers	187428	2423 (3.10-3.02)
RSRZ outliers	180081	2469 (3.10-3.02)

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	311	
1	B	311	
2	C	16	
2	E	16	
3	D	16	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	16	12%75%19%6%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5958 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 3-Methyladenine DNA glycosylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	289	Total	C	N	O	S	0	0	0
			2361	1528	382	444	7			
1	B	289	Total	C	N	O	S	0	0	0
			2357	1526	382	442	7			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	expression tag	UNP Q8R5T9
A	-12	ARG	-	expression tag	UNP Q8R5T9
A	-11	GLY	-	expression tag	UNP Q8R5T9
A	-10	SER	-	expression tag	UNP Q8R5T9
A	-9	HIS	-	expression tag	UNP Q8R5T9
A	-8	HIS	-	expression tag	UNP Q8R5T9
A	-7	HIS	-	expression tag	UNP Q8R5T9
A	-6	HIS	-	expression tag	UNP Q8R5T9
A	-5	HIS	-	expression tag	UNP Q8R5T9
A	-4	HIS	-	expression tag	UNP Q8R5T9
A	-3	GLY	-	expression tag	UNP Q8R5T9
A	-2	SER	-	expression tag	UNP Q8R5T9
A	-1	MET	-	expression tag	UNP Q8R5T9
A	0	ARG	-	expression tag	UNP Q8R5T9
B	-13	MET	-	expression tag	UNP Q8R5T9
B	-12	ARG	-	expression tag	UNP Q8R5T9
B	-11	GLY	-	expression tag	UNP Q8R5T9
B	-10	SER	-	expression tag	UNP Q8R5T9
B	-9	HIS	-	expression tag	UNP Q8R5T9
B	-8	HIS	-	expression tag	UNP Q8R5T9
B	-7	HIS	-	expression tag	UNP Q8R5T9
B	-6	HIS	-	expression tag	UNP Q8R5T9
B	-5	HIS	-	expression tag	UNP Q8R5T9
B	-4	HIS	-	expression tag	UNP Q8R5T9
B	-3	GLY	-	expression tag	UNP Q8R5T9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	SER	-	expression tag	UNP Q8R5T9
B	-1	MET	-	expression tag	UNP Q8R5T9
B	0	ARG	-	expression tag	UNP Q8R5T9

- Molecule 2 is a DNA chain called DNA (5'-D(*AP*GP*CP*GP*TP*CP*CP*AP*(3DR)P*GP*TP*CP*TP*AP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	16	Total	C	N	O	P	0	0	0
			311	149	54	93	15			
2	E	16	Total	C	N	O	P	0	0	0
			311	149	54	93	15			

- Molecule 3 is a DNA chain called DNA (5'-D(*T*GP*GP*TP*AP*GP*AP*CP*TP*TP*GP*GP*AP*CP*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	15	Total	C	N	O	P	0	0	0
			309	147	60	88	14			
3	F	15	Total	C	N	O	P	0	0	0
			309	147	60	88	14			

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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: 3-Methyladenine DNA glycosylase

Chain A: 



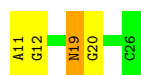
- Molecule 1: 3-Methyladenine DNA glycosylase

Chain B: 



- Molecule 2: DNA (5'-D(*AP*GP*CP*GP*TP*CP*CP*AP*(3DR)P*GP*TP*CP*TP*AP*CP*C)-3')

Chain C: 



- Molecule 2: DNA (5'-D(*AP*GP*CP*GP*TP*CP*CP*AP*(3DR)P*GP*TP*CP*TP*AP*CP*C)-3')

Chain E: 



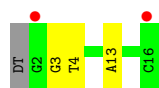
- Molecule 3: DNA (5'-D(*T*GP*GP*TP*AP*GP*AP*CP*TP*TP*GP*GP*AP*CP*GP*C)-3')

Chain D: 



- Molecule 3: DNA (5'-D(*T*GP*GP*TP*AP*GP*AP*CP*TP*TP*GP*GP*AP*CP*GP*C)-3')

Chain F:



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	56.80Å 103.60Å 66.72Å 90.00° 90.65° 90.00°	Depositor
Resolution (Å)	30.00 – 3.05 30.00 – 3.05	Depositor EDS
% Data completeness (in resolution range)	98.8 (30.00-3.05) 98.8 (30.00-3.05)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.74 (at 3.05Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.251 , 0.293 0.242 , 0.282	Depositor DCC
R_{free} test set	753 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	46.4	Xtriage
Anisotropy	0.869	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 0.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.034 for h,-k,-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	5958	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 35.53 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.7496e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 3DR

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	0.41	0/2414	0.75	0/3261
1	B	0.40	0/2410	0.74	0/3256
2	C	0.19	0/334	0.70	0/510
2	E	0.20	0/334	0.66	0/510
3	D	0.18	0/347	0.64	0/535
3	F	0.18	0/347	0.64	0/535
All	All	0.37	0/6186	0.73	0/8607

There are no bond length outliers.

There are no bond angle outliers.

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	2361	0	2332	6	0
1	B	2357	0	2328	6	0
2	C	311	0	178	3	0
2	E	311	0	178	8	0
3	D	309	0	170	7	0
3	F	309	0	170	6	0
All	All	5958	0	5356	34	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 3.

All (34) 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:13:DC:H2''	2:E:14:DG:C8	2.19	0.77
3:F:3:DG:C2'	3:F:4:DT:O5'	2.35	0.75
3:F:3:DG:H2'	3:F:4:DT:O5'	1.91	0.71
2:E:13:DC:H2''	2:E:14:DG:H8	1.62	0.65
3:F:3:DG:C5	3:F:4:DT:C4	2.86	0.64
2:E:18:DA:H3'	2:E:19:3DR:H5''	1.83	0.61
3:F:3:DG:C6	3:F:4:DT:N3	2.70	0.60
3:D:3:DG:H2''	3:D:4:DT:H5'	1.87	0.56
2:C:19:3DR:H2''	2:C:20:DG:OP2	2.06	0.56
3:D:4:DT:H2''	3:D:5:DA:N7	2.22	0.55
2:E:14:DG:H1	3:F:13:DA:N6	2.05	0.54
1:B:21:ASN:HD22	1:B:24:GLU:H	1.56	0.54
3:F:3:DG:C6	3:F:4:DT:C4	2.97	0.53
1:A:21:ASN:HD22	1:A:24:GLU:H	1.57	0.51
1:A:91:LEU:HB3	1:A:97:LEU:HD13	1.92	0.51
2:C:11:DA:H2'	2:C:12:DG:C8	2.46	0.51
2:E:25:DC:H1'	2:E:26:DC:C6	2.47	0.50
2:E:24:DA:H1'	2:E:25:DC:O5'	2.12	0.49
1:B:87:ILE:HG23	1:B:272:LEU:HD22	1.95	0.48
1:A:75:PHE:HA	1:A:80:LEU:HD12	1.95	0.48
1:B:91:LEU:HB3	1:B:97:LEU:HD13	1.96	0.48
3:D:3:DG:H2''	3:D:4:DT:C5'	2.44	0.48
3:D:15:DG:H2''	3:D:16:DC:C5'	2.44	0.47
1:B:75:PHE:HA	1:B:80:LEU:HD12	1.97	0.46
1:A:221:LYS:HG2	2:C:20:DG:H5'	1.97	0.45
3:D:7:DA:H1'	3:D:8:DC:H5'	1.99	0.45
2:E:23:DT:H2''	2:E:24:DA:C8	2.52	0.44
1:B:53:VAL:HG12	1:B:62:ILE:HG12	2.00	0.43
1:A:53:VAL:HG12	1:A:62:ILE:HG12	1.99	0.43
1:B:133:LYS:O	1:B:137:GLU:HG2	2.19	0.43
3:D:15:DG:H2''	3:D:16:DC:O5'	2.19	0.42
3:D:15:DG:H2'	3:D:16:DC:C6	2.55	0.42
2:E:14:DG:H5''	2:E:15:DT:OP2	2.20	0.42
1:A:15:ARG:HG2	1:A:59:THR:HG23	2.03	0.41

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	287/311 (92%)	282 (98%)	5 (2%)	0	100	100
1	B	287/311 (92%)	279 (97%)	7 (2%)	1 (0%)	36	63
All	All	574/622 (92%)	561 (98%)	12 (2%)	1 (0%)	43	70

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	254	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 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	255/275 (93%)	253 (99%)	2 (1%)	73	79
1	B	254/275 (92%)	253 (100%)	1 (0%)	84	84
All	All	509/550 (92%)	506 (99%)	3 (1%)	78	81

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	241	VAL
1	A	243	ILE
1	B	261	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	21	ASN
1	A	52	ASN
1	A	112	GLN
1	A	125	GLN
1	B	21	ASN
1	B	35	ASN
1	B	112	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	3DR	E	19	2	8,11,12	0.35	0	7,14,17	1.56	3 (42%)
2	3DR	C	19	2	8,11,12	0.41	0	7,14,17	1.68	3 (42%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	3DR	E	19	2	-	2/3/15/16	0/1/1/1
2	3DR	C	19	2	-	2/3/15/16	0/1/1/1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	19	3DR	C1'-O4'-C4'	2.56	112.40	108.37
2	C	19	3DR	C1'-C2'-C3'	2.41	105.84	103.26
2	E	19	3DR	C1'-C2'-C3'	2.35	105.78	103.26
2	E	19	3DR	O4'-C4'-C3'	2.32	107.14	103.73
2	E	19	3DR	C1'-O4'-C4'	2.20	111.83	108.37
2	C	19	3DR	O4'-C4'-C3'	2.09	106.81	103.73

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	19	3DR	C3'-C4'-C5'-O5'
2	C	19	3DR	C3'-C4'-C5'-O5'
2	E	19	3DR	O4'-C4'-C5'-O5'
2	C	19	3DR	O4'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	19	3DR	1	0
2	C	19	3DR	1	0

5.5 Carbohydrates [i](#)

There are no oligosaccharides 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	289/311 (92%)	-0.01	0 100 100	27, 45, 71, 76	0
1	B	289/311 (92%)	0.06	1 (0%) 90 80	31, 45, 77, 84	0
2	C	15/16 (93%)	0.58	0 100 100	44, 58, 79, 79	0
2	E	15/16 (93%)	1.03	1 (6%) 24 12	51, 72, 99, 100	0
3	D	15/16 (93%)	0.47	0 100 100	39, 57, 74, 75	0
3	F	15/16 (93%)	0.96	2 (13%) 7 4	46, 66, 95, 98	0
All	All	638/686 (93%)	0.09	4 (0%) 85 69	27, 46, 76, 100	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	11	DA	3.0
3	F	16	DC	2.6
1	B	286	GLU	2.3
3	F	2	DG	2.0

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	3DR	C	19	11/12	0.87	0.12	55,57,61,62	0
2	3DR	E	19	11/12	0.90	0.12	64,67,70,71	0

6.3 Carbohydrates [i](#)

There are no oligosaccharides 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.