



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 11, 2024 – 11:10 PM EDT

PDB ID : 1NGN
Title : Mismatch repair in methylated DNA. Structure of the mismatch-specific thymine glycosylase domain of methyl-CpG-binding protein MBD4
Authors : Wu, P.; Qiu, C.; Sohail, A.; Zhang, X.; Bhagwat, A.S.; Cheng, X.
Deposited on : 2002-12-17
Resolution : 2.10 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)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

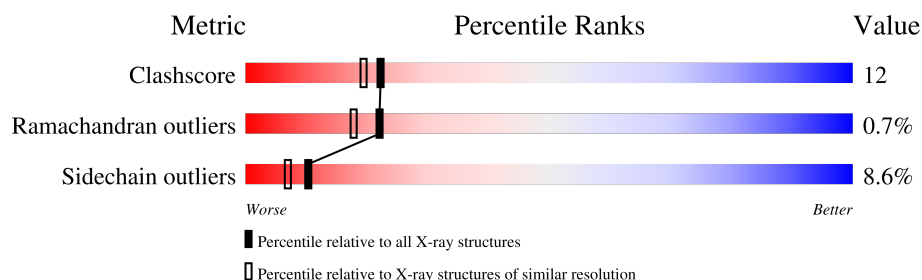
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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))
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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 $\leq 5\%$

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	155	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1304 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 methyl-CpG binding protein MBD4.

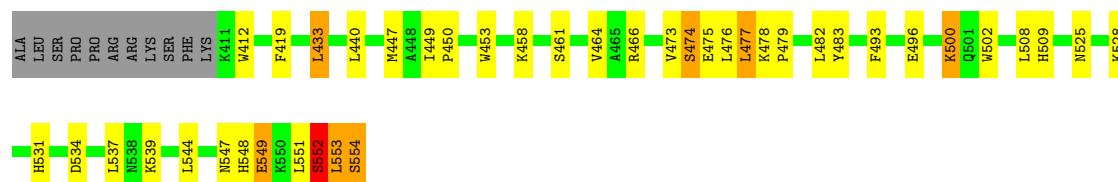
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	144	Total	C	N	O	S	0	1	0
			1214	797	204	211	2			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	90	Total	O	0	0
			90	90		

Note EDS was not executed.

- Molecule 1: methyl-CpG binding protein MBD4



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	48.58 Å 48.58 Å 146.57 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.33 – 2.10	Depositor
% Data completeness (in resolution range)	99.5 (19.33-2.10)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.213 , 0.263	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1304	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

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	0.39	0/1264	0.65	3/1720 (0.2%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	433	LEU	CA-CB-CG	6.63	130.54	115.30
1	A	553	LEU	CA-CB-CG	-5.34	103.01	115.30
1	A	553	LEU	N-CA-C	5.28	125.24	111.00

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	1214	0	1172	28	0
2	A	90	0	0	4	0
All	All	1304	0	1172	28	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 12.

The worst 5 of 28 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:531:HIS:HD2	2:A:92:HOH:O	1.79	0.65
1:A:528:LYS:HE3	1:A:554:SER:C	2.23	0.59
1:A:547:ASN:O	1:A:551:LEU:HG	2.03	0.57
1:A:458:LYS:HE2	1:A:476:LEU:HD12	1.85	0.57
1:A:551:LEU:O	1:A:552:SER:HB3	2.03	0.57

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	143/155 (92%)	140 (98%)	2 (1%)	1 (1%)	22	18

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	552	SER

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	129/142 (91%)	118 (92%)	11 (8%)	10	7

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	500	LYS
1	A	549	GLU
1	A	554	SER
1	A	552	SER
1	A	474	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	509	HIS
1	A	531	HIS
1	A	547	ASN
1	A	548	HIS

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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.