



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 12:58 PM GMT

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 full wwPDB validation report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

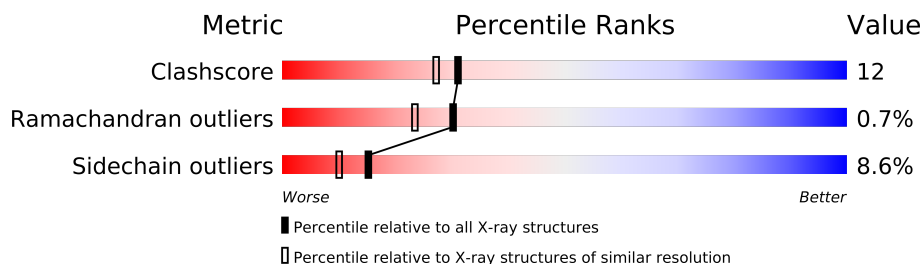
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 21963  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

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	79885	3649 (2.10-2.10)
Ramachandran outliers	78287	3610 (2.10-2.10)
Sidechain outliers	78261	3611 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

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 hydrogen and 0 are deuterium.

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		

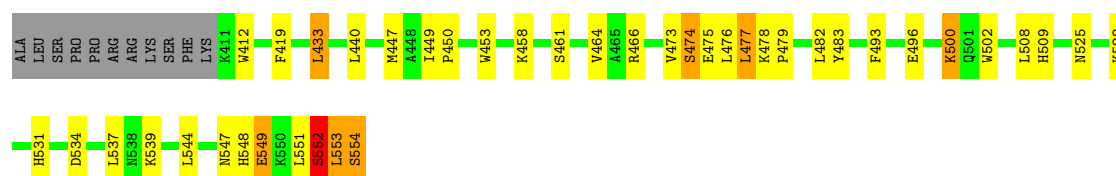
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

- Molecule 1: methyl-CpG binding protein MBD4

Chain A: 



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.58Å 48.58Å 146.57Å 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 <sup>2</sup> )	26.0	wwPDB-VP

## 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	0.39	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 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

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

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 12.

All (28) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(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:551:LEU:O	1:A:552:SER:HB3	2.03	0.57
1:A:458:LYS:HE2	1:A:476:LEU:HD12	1.85	0.57
1:A:412:TRP:CZ3	1:A:553:LEU:HD13	2.40	0.57
1:A:525:ASN:HA	2:A:89:HOH:O	2.05	0.56
1:A:549:GLU:HB2	2:A:30:HOH:O	2.07	0.53
1:A:419:PHE:CD2	1:A:539:LYS:HE3	2.42	0.53
1:A:502:TRP:CD1	1:A:508:LEU:HD21	2.44	0.53
1:A:544:LEU:O	1:A:548:HIS:HB3	2.09	0.52
1:A:461:SER:OG	1:A:464:VAL:HG23	2.09	0.52
1:A:475:GLU:OE1	1:A:478:LYS:HE3	2.08	0.52
1:A:496:GLU:O	1:A:500:LYS:HB2	2.12	0.50
1:A:548:HIS:HD2	1:A:554:SER:O	1.96	0.49
1:A:493:PHE:HB2	1:A:509:HIS:O	2.13	0.48
1:A:474:SER:HB2	1:A:483:TYR:O	2.13	0.48
1:A:500:LYS:HE2	1:A:500:LYS:HB3	1.65	0.47
1:A:534:ASP:HB3	1:A:537:LEU:HB2	1.97	0.46
1:A:419:PHE:CE2	1:A:539:LYS:HE3	2.51	0.46
1:A:473:VAL:HG12	1:A:477:LEU:HD22	1.98	0.46
1:A:412:TRP:CH2	1:A:553:LEU:HD13	2.50	0.46
1:A:478:LYS:N	1:A:479:PRO:HD2	2.32	0.44
1:A:453:TRP:HA	1:A:453:TRP:CE3	2.52	0.43
1:A:478:LYS:N	1:A:479:PRO:CD	2.83	0.42
1:A:412:TRP:HZ3	1:A:553:LEU:HD22	1.85	0.41
1:A:500:LYS:NZ	2:A:86:HOH:O	2.43	0.41
1:A:449:ILE:HB	1:A:450:PRO:HD3	2.03	0.40

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	143/155 (92%)	140 (98%)	2 (1%)	1 (1%)	30	23

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	552	SER

### 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	129/142 (91%)	118 (92%)	11 (8%)	15	10

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	433	LEU
1	A	440	LEU
1	A	447	MET
1	A	466	ARG
1	A	474	SER
1	A	477	LEU
1	A	482	LEU
1	A	500	LYS
1	A	549	GLU
1	A	552	SER
1	A	554	SER

Some 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 ⓘ

There are no RNA chains in this entry.



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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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 will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.