



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 02:45 AM UTC

PDB ID : 6YWW / pdb_00006yww
Title : MeCP2 is a microsatellite binding protein that protects CA repeats from nucleosome invasion
Authors : Ibrahim, A.; Papin, C.; Mohideen-Abdul, K.; Gras, S.L.; Stoll, I.; Bronner, C.; Dimitrov, S.; Klaholz, B.P.; Hamiche, A.
Deposited on : 2020-04-30
Resolution : 2.10 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

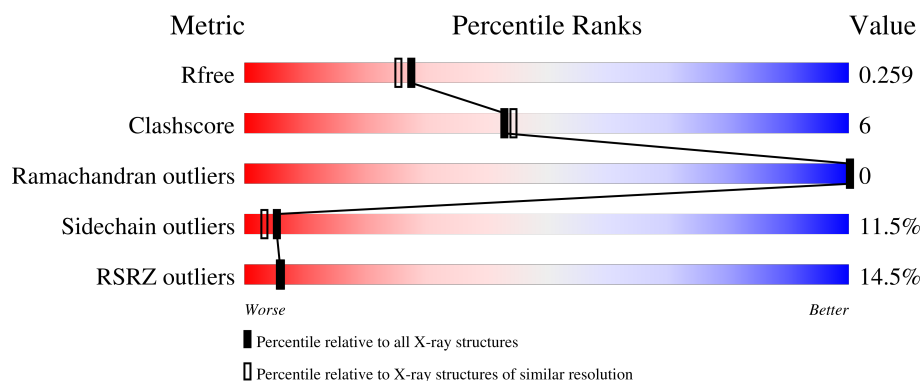
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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))
R_{free}	180053	6658 (2.10-2.10)
Clashscore	190562	7164 (2.10-2.10)
Ramachandran outliers	187476	7099 (2.10-2.10)
Sidechain outliers	187428	7100 (2.10-2.10)
RSRZ outliers	180081	6662 (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 $\leq 5\%$. 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	101	16% 54% 13% • 30%
2	B	20	65% 35%
3	C	20	80% 20%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 1468 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 Truncated methyl CpG binding protein 2 transcript 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	71	Total	C	N	O	S	0	0	0
			570	363	98	108	1			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	76	MET	-	initiating methionine	UNP D5L9I4
A	162	ARG	-	expression tag	UNP D5L9I4
A	163	GLY	-	expression tag	UNP D5L9I4
A	164	ARG	-	expression tag	UNP D5L9I4
A	165	GLY	-	expression tag	UNP D5L9I4
A	166	SER	-	expression tag	UNP D5L9I4
A	167	PRO	-	expression tag	UNP D5L9I4
A	168	ALA	-	expression tag	UNP D5L9I4
A	169	ALA	-	expression tag	UNP D5L9I4
A	170	ALA	-	expression tag	UNP D5L9I4
A	171	HIS	-	expression tag	UNP D5L9I4
A	172	HIS	-	expression tag	UNP D5L9I4
A	173	HIS	-	expression tag	UNP D5L9I4
A	174	HIS	-	expression tag	UNP D5L9I4
A	175	HIS	-	expression tag	UNP D5L9I4
A	176	HIS	-	expression tag	UNP D5L9I4

- Molecule 2 is a DNA chain called DNA/RNA (5'-D(*TP*CP*TP*GP*CP*AP*CP*A)-R(P*(5HC))-D(P*AP*CP*AP*CP*AP*AP*TP*TP*AP*TP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	20	Total	C	N	O	P	0	0	0
			403	195	73	116	19			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	20	Total 413	C 199	N 74	O 121	P 19	0	0	0

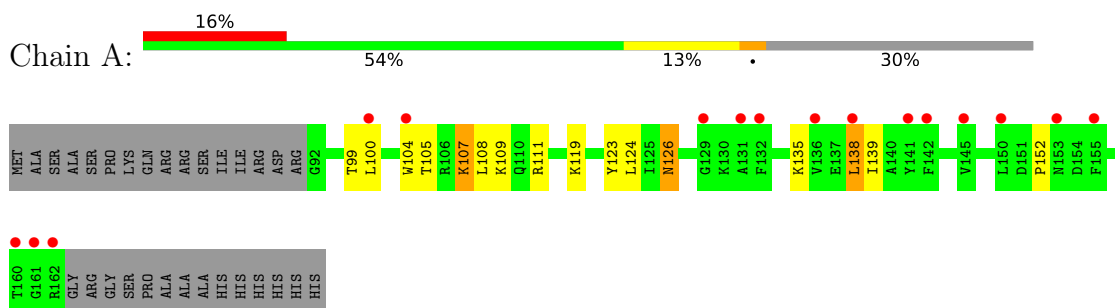
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	24	Total 24	O 24	0	0
4	B	31	Total 31	O 31	0	0
4	C	27	Total 27	O 27	0	0

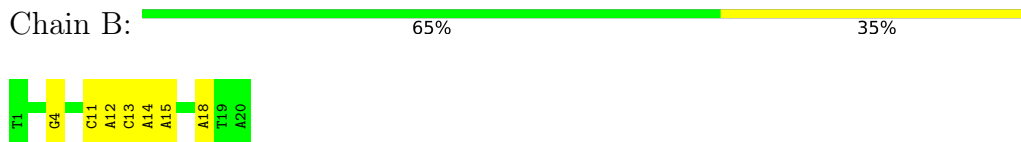
3 Residue-property plots [i](#)

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.

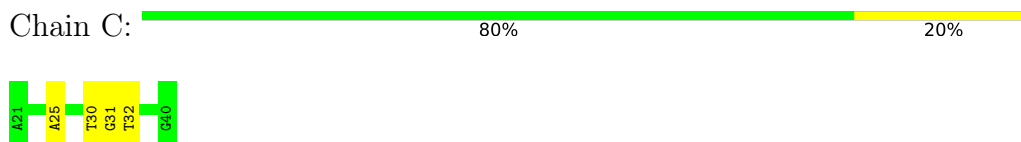
- Molecule 1: Truncated methyl CpG binding protein 2 transcript 1



- Molecule 2: DNA/RNA (5'-D(*TP*CP*TP*GP*CP*AP*CP*A)-R(P*(5HC))-D(P*AP*CP*A P*CP*AP*AP*TP*TP*AP*TP*A)-3')



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



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	64.38Å 48.99Å 66.19Å 90.00° 90.53° 90.00°	Depositor
Resolution (Å)	66.18 – 2.10 66.18 – 2.10	Depositor EDS
% Data completeness (in resolution range)	50.9 (66.18-2.10) 50.9 (66.18-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.39 (at 2.10Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.221 , 0.259 0.232 , 0.259	Depositor DCC
R_{free} test set	312 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	46.5	Xtriage
Anisotropy	0.209	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 51.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.036 for -h,-k,l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	1468	wwPDB-VP
Average B, all atoms (Å ²)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.31% 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

Bond lengths and bond angles in the following residue types are not validated in this section: 5HC

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	0.78	0/584	1.07	1/787 (0.1%)
2	B	0.74	3/427 (0.7%)	0.86	2/653 (0.3%)
3	C	0.46	0/463	0.72	0/715
All	All	0.68	3/1474 (0.2%)	0.90	3/2155 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	15	DA	O3'-P	-5.85	1.52	1.61
2	B	13	DC	O3'-P	-5.66	1.52	1.61
2	B	14	DA	O3'-P	-5.32	1.53	1.61

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	126	ASN	CA-CB-CG	7.64	120.24	112.60
2	B	15	DA	C2'-C3'-O3'	-7.26	100.61	111.50
2	B	13	DC	C2'-C3'-O3'	-6.42	101.86	111.50

There are no chirality outliers.

There are no planarity outliers.

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	570	0	559	9	0
2	B	403	0	229	3	0
3	C	413	0	230	3	0
4	A	24	0	0	0	0
4	B	31	0	0	1	0
4	C	27	0	0	0	0
All	All	1468	0	1018	14	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.

All (14) 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)
3:C:30:DT:H2''	3:C:31:DG:O5'	1.86	0.76
1:A:104:TRP:CZ3	1:A:126:ASN:HB3	2.29	0.68
1:A:104:TRP:CZ3	1:A:126:ASN:CB	2.81	0.64
1:A:104:TRP:CE3	1:A:126:ASN:HA	2.34	0.63
1:A:111:ARG:HG3	1:A:119:LYS:O	2.10	0.51
1:A:135:LYS:HA	1:A:138:LEU:HD12	1.92	0.50
1:A:139:ILE:HG12	1:A:152:PRO:HG2	1.96	0.48
3:C:31:DG:H2'	3:C:32:DT:H71	1.97	0.47
2:B:18:DA:C2	3:C:25:DA:C2	3.05	0.44
1:A:107:LYS:HB3	1:A:123:TYR:HB2	2.00	0.44
1:A:109:LYS:HB2	1:A:123:TYR:HE1	1.82	0.44
1:A:104:TRP:CH2	1:A:126:ASN:HB2	2.53	0.43
2:B:4:DG:OP2	4:B:101:HOH:O	2.21	0.42
2:B:11:DC:H2'	2:B:12:DA:C8	2.55	0.42

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	69/101 (68%)	63 (91%)	6 (9%)	0	100	100

There are no Ramachandran outliers to report.

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	61/84 (73%)	54 (88%)	7 (12%)	5	3

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	99	THR
1	A	100	LEU
1	A	105	THR
1	A	107	LYS
1	A	108	LEU
1	A	124	LEU
1	A	138	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	5HC	B	9	3,2	18,22,23	0.52	0	22,31,34	0.35	0

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	5HC	B	9	3,2	-	5/9/23/24	0/2/2/2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	9	5HC	C3'-C4'-C5'-O5'
2	B	9	5HC	O4'-C4'-C5'-O5'
2	B	9	5HC	C4-C5-C5M-O5
2	B	9	5HC	C6-C5-C5M-O5
2	B	9	5HC	C4'-C5'-O5'-P

There are no ring outliers.

No monomer is involved in short contacts.

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

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 [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	71/101 (70%)	1.18	16 (22%) 2 2	29, 47, 68, 75	0
2	B	19/20 (95%)	0.57	0 100 100	39, 44, 52, 55	0
3	C	20/20 (100%)	0.62	0 100 100	41, 50, 61, 62	0
All	All	110/141 (78%)	0.97	16 (14%) 6 6	29, 48, 66, 75	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	141	TYR	4.0
1	A	145	VAL	3.7
1	A	104	TRP	3.1
1	A	161	GLY	3.0
1	A	100	LEU	2.8
1	A	142	PHE	2.7
1	A	162	ARG	2.5
1	A	155	PHE	2.5
1	A	131	ALA	2.4
1	A	138	LEU	2.4
1	A	129	GLY	2.4
1	A	153	ASN	2.2
1	A	150	LEU	2.2
1	A	160	THR	2.2
1	A	136	VAL	2.2
1	A	132	PHE	2.1

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($\text{\AA}^2$)	Q<0.9
2	5HC	B	9	21/22	0.93	0.09	37,41,52,55	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.