



wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2020 – 09:48 am BST

PDB ID : 3AV6
Title : Crystal structure of mouse DNA methyltransferase 1 with AdoMet
Authors : Takeshita, K.; Suetake, I.; Yamashita, E.; Suga, M.; Narita, H.; Nakagawa, A.; Tajima, S.
Deposited on : 2011-02-22
Resolution : 3.09 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
buster-report : 1.1.7 (2018)
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) : 2.11

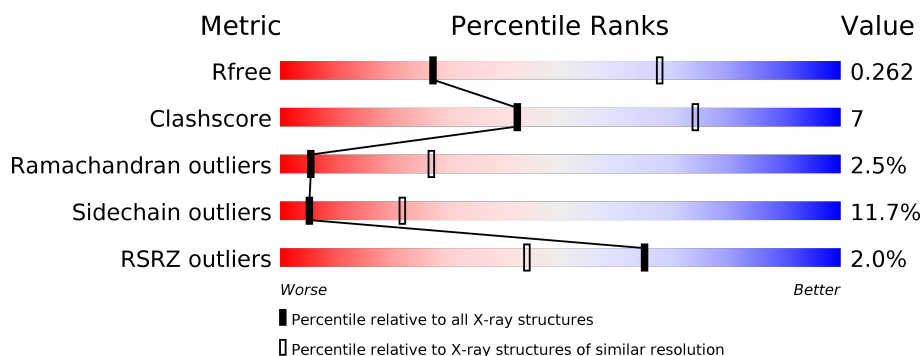
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.09 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 ≥ 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	1330	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9110 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.

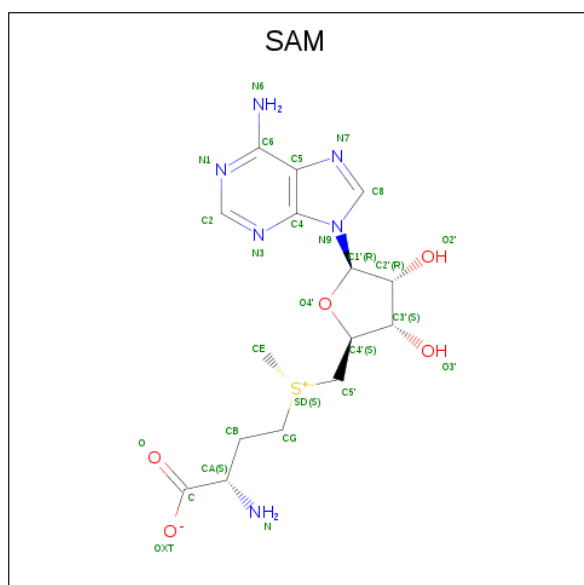
- Molecule 1 is a protein called DNA (cytosine-5)-methyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1135	Total	C	N	O	S	0	0	0
			9079	5744	1579	1697	59			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	4	Total	Zn	0	0
			4	4		

- Molecule 3 is S-ADENOSYLMETHIONINE (three-letter code: SAM) (formula: C₁₅H₂₂N₆O₅S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			27	15	6	5	1		

H1543	W1439	Q1306
H1544	R1440	T1312
E1545		R1313
Q1546	Q1446	R1314
V1549	V1447	I1317
R1553	R1448	
	L1449	A1320
R1557	G1450	F1330
	D1451	P1333
	G1452	V1336
H1575	A1455	F1337
R1576	L1458	V1346
Q1577	Q1459	V1354
A1587		S1355
S1601	F1462	N1356
	H1463	R1359
A1603	D1464	L1360
ALA	L1475	S1361
VAL		S1362
LYS	V1478	R1366
ALA	C1479	V1370
LYS	S1480	R1371
GLU	G1481	
GLU	A1482	M1374
ALA		S1385
ALA	D1488	E1394
THR	P1489	P1395
LYS	E1490	L1403
ASP	S1491	H1407
	R1492	Y1408
	W1500	Q1409
		P1410
	H1504	I1411
	N1507	L1412
	R1508	R1413
	A1513	T1416
	Y1516	C1417
	D1522	K1418
	F1525	D1419
	T1528	M1420
	V1529	A1426
	T1530	H1430
	N1531	
	P1532	
	K1537	
	V1541	
	L1542	

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	135.00Å 97.14Å 130.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.14 – 3.09 43.14 – 3.09	Depositor EDS
% Data completeness (in resolution range)	(Not available) (43.14-3.09) 99.6 (43.14-3.09)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	6.02 (at 3.12Å)	Xtriage
Refinement program	BUSTER-TNT, BUSTER 2.8.0	Depositor
R, R_{free}	0.191 , 0.255 0.198 , 0.262	Depositor DCC
R_{free} test set	1626 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	54.2	Xtriage
Anisotropy	0.165	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 64.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.016 for l,-k,h	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	9110	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.39% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, SAM

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	0.52	0/9299	0.79	2/12573 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	1360	LEU	C-N-CA	7.21	139.73	121.70
1	A	1042	ILE	C-N-CA	6.25	137.34	121.70

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	9079	0	8847	124	0
2	A	4	0	0	0	0
3	A	27	0	22	0	0
All	All	9110	0	8869	124	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 7.

The worst 5 of 124 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:1543:HIS:HD2	1:A:1546:GLN:H	1.21	0.88
1:A:885:LYS:HA	1:A:886:THR:HG23	1.60	0.82
1:A:975:GLU:OE1	1:A:978:ARG:HD2	1.81	0.80
1:A:886:THR:H	1:A:887:GLN:HG2	1.50	0.77
1:A:1543:HIS:CD2	1:A:1546:GLN:H	2.02	0.76

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	1113/1330 (84%)	991 (89%)	94 (8%)	28 (2%)	5	27

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	465	GLY
1	A	573	ALA
1	A	772	ASP
1	A	777	LEU
1	A	896	PHE

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	1000/1162 (86%)	883 (88%)	117 (12%)	5 22

5 of 117 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	872	TRP
1	A	990	ASP
1	A	1480	SER
1	A	882	SER
1	A	917	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1076	GLN
1	A	1159	HIS
1	A	1430	HIS
1	A	953	ASN
1	A	1446	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 ⓘ

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

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	$\# Z > 2$
3	SAM	A	1	-	21,29,29	1.17	0	18,42,42	2.29	6 (33%)

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
3	SAM	A	1	-	-	3/8/33/33	0/3/3/3

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	A	1	SAM	O3'-C3'-C4'	-4.88	96.94	111.05
3	A	1	SAM	N3-C2-N1	-4.04	122.36	128.68
3	A	1	SAM	C5'-SD-CG	3.72	112.90	103.40
3	A	1	SAM	C4-C5-N7	-3.08	106.19	109.40
3	A	1	SAM	O4'-C1'-C2'	-3.05	102.47	106.93

There are no chirality outliers.

All (3) torsion outliers are listed below:

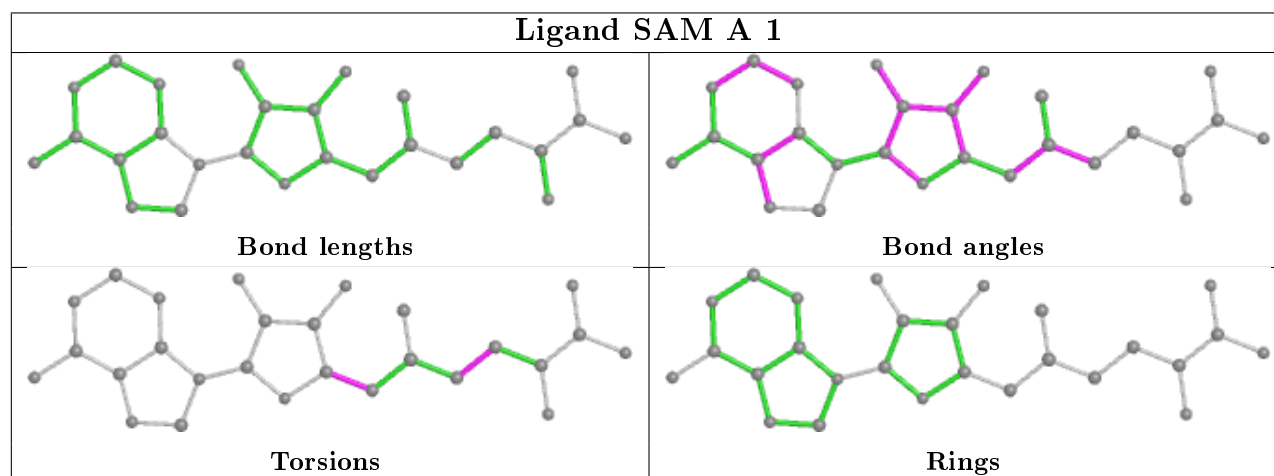
Mol	Chain	Res	Type	Atoms
3	A	1	SAM	CA-CB-CG-SD
3	A	1	SAM	O4'-C4'-C5'-SD
3	A	1	SAM	C3'-C4'-C5'-SD

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be

highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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	1135/1330 (85%)	-0.13	23 (2%) 65 44	13, 52, 100, 127	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	408	SER	4.3
1	A	390	GLU	4.2
1	A	923	GLY	3.9
1	A	379	ASP	3.2
1	A	378	GLU	3.1

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [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.

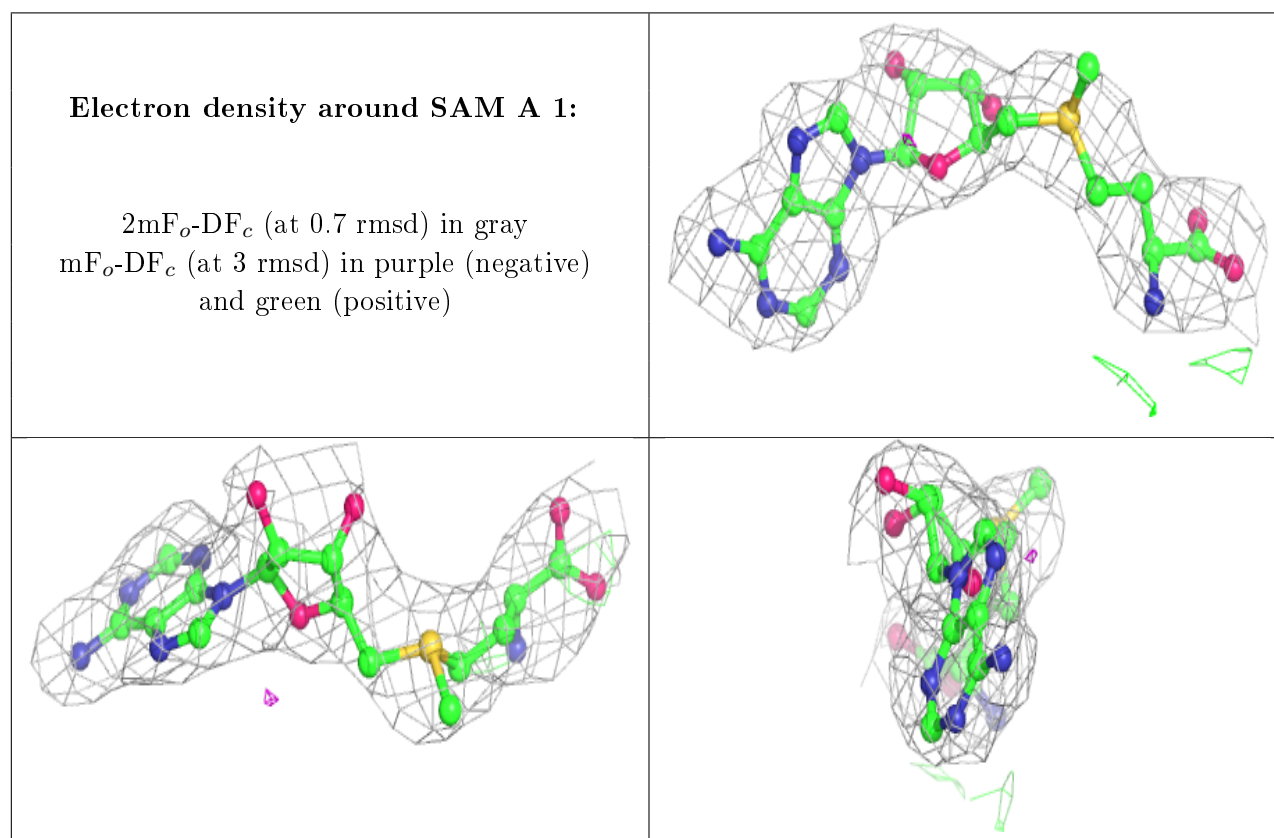
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	SAM	A	1	27/27	0.97	0.17	42,43,50,51	0
2	ZN	A	2001	1/1	0.99	0.14	55,55,55,55	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	ZN	A	2005	1/1	1.00	0.11	33,33,33,33	0
2	ZN	A	2002	1/1	1.00	0.07	58,58,58,58	0
2	ZN	A	2004	1/1	1.00	0.10	47,47,47,47	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



6.5 Other polymers [i](#)

There are no such residues in this entry.