



wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 09:02 PM GMT

PDB ID : 2PV0
Title : DNA methyltransferase 3 like protein (DNMT3L)
Authors : Cheng, X.
Deposited on : 2007-05-09
Resolution : 3.30 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

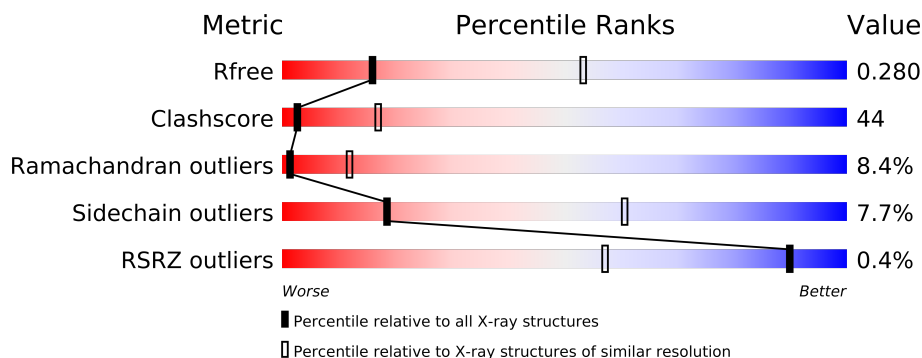
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
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 3.30 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}	66092	1341 (3.40-3.20)
Clashscore	79885	1696 (3.40-3.20)
Ramachandran outliers	78287	1664 (3.40-3.20)
Sidechain outliers	78261	1662 (3.40-3.20)
RSRZ outliers	66119	1342 (3.40-3.20)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	386	
1	B	386	
1	C	386	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	ZN	C	508	-	X

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7861 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 DNA (cytosine-5)-methyltransferase3-like.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	347	Total	C	N	O	S	0	0	0
			2737	1748	463	505	21			
1	A	347	Total	C	N	O	S	0	0	0
			2744	1753	467	503	21			
1	C	347	Total	C	N	O	S	0	0	0
			2371	1478	415	459	19			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	278	GLY	ARG	CONFLICT	UNP Q9UJW3
B	?	-	SER	DELETION	UNP Q9UJW3
A	278	GLY	ARG	CONFLICT	UNP Q9UJW3
A	?	-	SER	DELETION	UNP Q9UJW3
C	278	GLY	ARG	CONFLICT	UNP Q9UJW3
C	?	-	SER	DELETION	UNP Q9UJW3

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

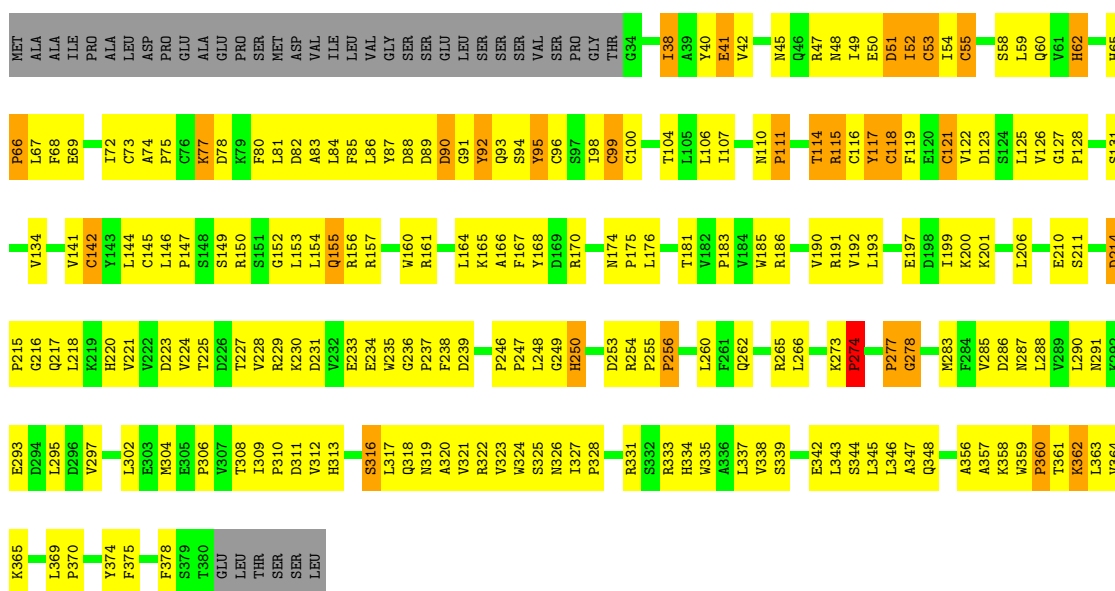
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	3	Total	Zn	0	0
			3	3		
2	A	3	Total	Zn	0	0
			3	3		
2	C	3	Total	Zn	0	0
			3	3		

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.

- Molecule 1: DNA (cytosine-5)-methyltransferase3-like

Chain B:



Y374
F375
K376
Y377
F378
S379
T380
GLU
LEU
THR
SER
SER
LEU

● Molecule 1: DNA (cytosine-5)-methyltransferase3-like

Chain C: 

MET	ALA	ALA	ILE	PRO	ALA	LEU	ASP	PRO	GLU	ALA	GLU	SER	MET	ASP	VAL	ILE	LEU	VAL	GLY	SER	SER	GLU	LEU	SER	SER	SER	VAL	SER	PRO	PRO	GLY	THR	G34	R35	D36	L37	I38	E41	V42	M45	Q46	R47	M48	I49	E50	D51	I52	C53	I54	C55	C56	G57	S58	L59	H62	T63
Q64	H65	P66	L67	F68	E69	G70	G71	I72	C73	A74	P75	A83	L84	F85	L86	Y87	D88	D89	D90	G91	Y92	Q93	S94	Y95	C96	S97	I98	C99	C100	E103	T104	L105	L106	I107	C108	G109	M110	P111	D112	Y117	C118	F119	E120	C121	V122	D123	S124	L125	V126	G127	P128	G129	T130	V134	T63	
M137	S138	M139	W140	V141	C142	Y143	L144	C145	L153	L154	Q155	R156	R157	R158	K159	W160	R161	S162	Q163	Y168	D169	S172	E173	M174	P175	L176	E177	M178	F179	E180	T181	V182	P183	R187	Q188	P189	V190	R191	V192	L193	S194	F196	E197	D198	I199	K200	K201	E202	P203	E210	D214	P215				
G216	Q217	L218	K219	H220	V224	T225	D226	T227	V228	R229	K230	D231	V232	G236	P237	F238	D239	L240	V241	Y242	G243	L244	T245	P246	P247	L248	G249	H250	T251	C252	P255	P256	L260	F261	Q262	F263	H264	R265	L266	L267	Q268	R271	P272	K273	P274	G275	S276	P277	G278	P279	F280	W281	W282	M283		
F284	V285	D286	N287	E293	D294	L295	D296	V297	A298	S299	R300	E303	P306	I309	F310	D311	V312	H313	G314	G315	S316	L317	A320	V323	W324	S325	N326	I327	P328	A329	I330	W335	A336	L337	V338	S339	E342	L343	S344	L345	L346	A347	Q348	N349	K350	Q351	A356	A357	K358	W359						
P360	T361	K362	L363	V364	K365	N366	C367	F368	L369	P370	L371	R372	E373	Y374	F375	K376	Y377	F378	S379	T380	GLU	LEU	THR	SER	SER	LEU																														

4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α , β , γ	267.20Å 267.20Å 149.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.74 – 3.30 34.37 – 3.29	Depositor EDS
% Data completeness (in resolution range)	98.8 (29.74-3.30) 98.8 (34.37-3.29)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.72 (at 3.32Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.248 , 0.272 0.263 , 0.280	Depositor DCC
R_{free} test set	2344 reflections (4.99%)	DCC
Wilson B-factor (Å ²)	108.2	Xtriage
Anisotropy	0.020	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 70.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.53$, $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	0 of 47265 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7861	wwPDB-VP
Average B, all atoms (Å ²)	113.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.80% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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

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	0.42	0/2821	0.68	1/3840 (0.0%)
1	B	0.42	0/2814	0.69	0/3833
1	C	0.31	0/2431	0.57	0/3348
All	All	0.39	0/8066	0.65	1/11021 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	369	LEU	CA-CB-CG	-5.19	103.37	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	374	TYR	Sidechain

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	2744	0	2625	212	0
1	B	2737	0	2605	237	0
1	C	2371	0	1927	217	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
2	C	3	0	0	0	0
All	All	7861	0	7157	659	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 44.

The worst 5 of 659 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:118:CYS:HB2	1:A:121:CYS:HB2	1.18	1.13
1:B:155:GLN:HA	1:B:155:GLN:HE21	1.17	1.06
1:C:47:ARG:HB3	1:C:47:ARG:HH11	1.15	1.05
1:B:65:HIS:ND1	1:B:72:ILE:HD11	1.72	1.04
1:C:118:CYS:HB2	1:C:121:CYS:HB2	1.38	1.03

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	345/386 (89%)	268 (78%)	53 (15%)	24 (7%)	2 17
1	B	345/386 (89%)	254 (74%)	66 (19%)	25 (7%)	2 16
1	C	345/386 (89%)	226 (66%)	81 (24%)	38 (11%)	1 6
All	All	1035/1158 (89%)	748 (72%)	200 (19%)	87 (8%)	1 12

5 of 87 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	89	ASP
1	B	142	CYS
1	B	147	PRO
1	B	235	TRP
1	B	277	PRO

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	301/344 (88%)	279 (93%)	22 (7%)	20	63
1	B	300/344 (87%)	278 (93%)	22 (7%)	20	63
1	C	208/344 (60%)	190 (91%)	18 (9%)	15	53
All	All	809/1032 (78%)	747 (92%)	62 (8%)	18	60

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	183	PRO
1	A	271	ARG
1	C	168	TYR
1	A	242	TYR
1	A	294	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	155	GLN
1	A	268	GLN
1	C	264	HIS
1	A	174	ASN
1	A	264	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 ⓘ

Of 9 ligands modelled in this entry, 9 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 ⓘ

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	347/386 (89%)	-0.21	0	100 100	29, 85, 169, 195	0
1	B	347/386 (89%)	-0.19	0	100 100	33, 81, 191, 200	0
1	C	347/386 (89%)	0.11	4 (1%)	75 29	93, 166, 199, 200	0
All	All	1041/1158 (89%)	-0.10	4 (0%)	90 57	29, 112, 195, 200	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	320	ALA	4.4
1	C	250	HIS	3.6
1	C	252	CYS	2.5
1	C	376	LYS	2.4

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	ZN	C	508	1/1	0.29	2.66	104,104,104,104	1
2	ZN	B	505	1/1	0.24	1.12	13,13,13,13	1
2	ZN	A	502	1/1	0.23	0.74	48,48,48,48	1
2	ZN	B	506	1/1	0.17	0.07	67,67,67,67	0
2	ZN	A	503	1/1	0.15	-0.05	75,75,75,75	0
2	ZN	B	504	1/1	0.13	-0.19	59,59,59,59	1
2	ZN	A	501	1/1	0.11	-1.06	91,91,91,91	1
2	ZN	C	509	1/1	0.06	-1.43	127,127,127,127	0
2	ZN	C	507	1/1	0.03	-2.40	167,167,167,167	1

6.5 Other polymers ⓘ

There are no such residues in this entry.