



wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 06:57 AM GMT

PDB ID : 2QRV
Title : Structure of Dnmt3a-Dnmt3L C-terminal domain complex
Authors : Jia, D.; Cheng, X.
Deposited on : 2007-07-29
Resolution : 2.89 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>

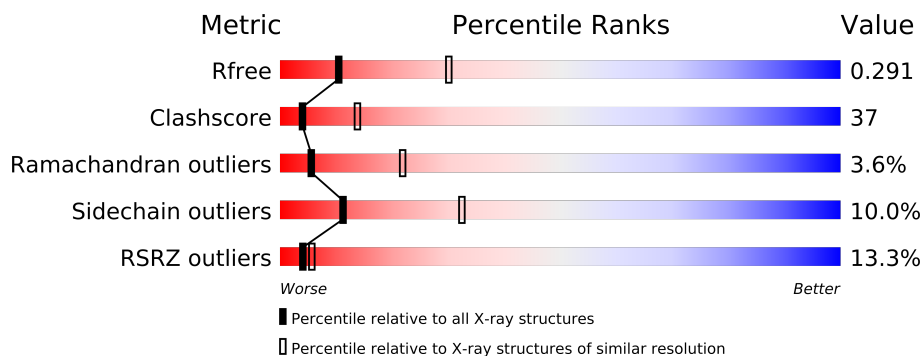
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.89 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	1053 (2.90-2.90)
Clashscore	79885	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RSRZ outliers	66119	1054 (2.90-2.90)

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	295	
1	D	295	
1	E	295	
1	H	295	
2	B	230	
2	C	230	
2	F	230	
2	G	230	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 14657 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)-methyltransferase3A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	272	Total	C	N	O	S	0	0	0
			2143	1370	379	381	13			
1	D	267	Total	C	N	O	S	0	0	0
			2111	1350	372	376	13			
1	E	270	Total	C	N	O	S	0	0	0
			2128	1361	375	379	13			
1	H	267	Total	C	N	O	S	0	0	0
			2111	1350	372	376	13			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	614	MET	-	EXPRESSION TAG	UNP Q9Y6K1
A	615	GLY	-	EXPRESSION TAG	UNP Q9Y6K1
A	616	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
A	617	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
A	618	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
A	619	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
A	620	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
A	621	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
A	622	MET	-	EXPRESSION TAG	UNP Q9Y6K1
D	614	MET	-	EXPRESSION TAG	UNP Q9Y6K1
D	615	GLY	-	EXPRESSION TAG	UNP Q9Y6K1
D	616	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
D	617	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
D	618	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
D	619	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
D	620	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
D	621	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
D	622	MET	-	EXPRESSION TAG	UNP Q9Y6K1
E	614	MET	-	EXPRESSION TAG	UNP Q9Y6K1
E	615	GLY	-	EXPRESSION TAG	UNP Q9Y6K1
E	616	HIS	-	EXPRESSION TAG	UNP Q9Y6K1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	617	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
E	618	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
E	619	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
E	620	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
E	621	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
E	622	MET	-	EXPRESSION TAG	UNP Q9Y6K1
H	614	MET	-	EXPRESSION TAG	UNP Q9Y6K1
H	615	GLY	-	EXPRESSION TAG	UNP Q9Y6K1
H	616	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
H	617	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
H	618	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
H	619	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
H	620	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
H	621	HIS	-	EXPRESSION TAG	UNP Q9Y6K1
H	622	MET	-	EXPRESSION TAG	UNP Q9Y6K1

- Molecule 2 is a protein called DNA (cytosine-5)-methyltransferase3-like.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	186	Total	C	N	O	S	0	0	0
			1515	991	254	266	4			
2	C	186	Total	C	N	O	S	0	0	0
			1515	991	254	266	4			
2	F	186	Total	C	N	O	S	0	0	0
			1515	991	254	266	4			
2	G	186	Total	C	N	O	S	0	0	0
			1515	991	254	266	4			

There are 12 discrepancies between the modelled and reference sequences:

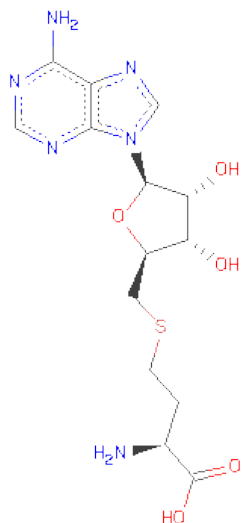
Chain	Residue	Modelled	Actual	Comment	Reference
B	157	GLY	-	EXPRESSION TAG	UNP Q9UJW3
B	158	SER	-	EXPRESSION TAG	UNP Q9UJW3
B	159	MET	-	EXPRESSION TAG	UNP Q9UJW3
C	157	GLY	-	EXPRESSION TAG	UNP Q9UJW3
C	158	SER	-	EXPRESSION TAG	UNP Q9UJW3
C	159	MET	-	EXPRESSION TAG	UNP Q9UJW3
F	157	GLY	-	EXPRESSION TAG	UNP Q9UJW3
F	158	SER	-	EXPRESSION TAG	UNP Q9UJW3
F	159	MET	-	EXPRESSION TAG	UNP Q9UJW3
G	157	GLY	-	EXPRESSION TAG	UNP Q9UJW3
G	158	SER	-	EXPRESSION TAG	UNP Q9UJW3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	159	MET	-	EXPRESSION TAG	UNP Q9UJW3

- Molecule 3 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: $C_{14}H_{20}N_6O_5S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
3	D	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
3	E	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
3	H	1	Total	C	N	O	S	0	0
			26	14	6	5	1		

A347	Q348	N349	K350	Q351	S352	S353	LYS	LEU	ALA	ALA	LYS	THR	PRO	T361	K362	L363	V364	K365	N366	C367	F368	L369	P370	L371	R372	E373	Y374	F375	K376	Y377	F378	S379	THR	GLU	LEU	THR	SER	SER	LEU
------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----

4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	401.88Å 401.88Å 49.69Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.90 – 2.89 43.85 – 2.90	Depositor EDS
% Data completeness (in resolution range)	81.9 (38.90-2.89) 92.7 (43.85-2.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.32 (at 2.90Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.259 , 0.281 0.278 , 0.291	Depositor DCC
R_{free} test set	3124 reflections (5.07%)	DCC
Wilson B-factor (Å ²)	77.8	Xtriage
Anisotropy	0.423	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 57.5	EDS
Estimated twinning fraction	0.014 for h,-h-k,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 65264 reflections	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	14657	wwPDB-VP
Average B, all atoms (Å ²)	101.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% 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: SAH

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	0.55	0/2196	0.83	2/2973 (0.1%)
1	D	0.52	0/2162	0.76	2/2926 (0.1%)
1	E	0.58	0/2180	0.78	1/2951 (0.0%)
1	H	0.44	0/2162	0.65	0/2926
2	B	0.41	0/1560	0.61	0/2120
2	C	0.37	0/1560	0.59	0/2120
2	F	0.35	0/1560	0.58	0/2120
2	G	0.36	0/1560	0.54	0/2120
All	All	0.47	0/14940	0.69	5/20256 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	844	PHE	C-N-CD	-13.46	91.00	120.60
1	D	875	ASN	N-CA-C	-6.49	93.47	111.00
1	A	644	LEU	CA-CB-CG	6.20	129.55	115.30
1	D	872	ASP	CB-CG-OD2	-5.62	113.24	118.30
1	E	881	ARG	NE-CZ-NH1	5.04	122.82	120.30

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	2143	0	2076	174	0
1	D	2111	0	2052	172	0
1	E	2128	0	2067	170	0
1	H	2111	0	2052	195	0
2	B	1515	0	1484	122	0
2	C	1515	0	1484	93	0
2	F	1515	0	1484	109	0
2	G	1515	0	1484	105	0
3	A	26	0	19	1	0
3	D	26	0	19	0	0
3	E	26	0	19	1	0
3	H	26	0	19	2	0
All	All	14657	0	14259	1076	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 37.

The worst 5 of 1076 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:E:786:ARG:HG3	1:E:786:ARG:HH11	1.19	1.06
1:E:655:ARG:HH12	1:E:695:GLY:HA3	1.20	1.05
2:B:271:ARG:HH11	2:B:271:ARG:HG3	1.13	1.05
1:E:854:ILE:HG22	1:E:855:LEU:H	1.18	1.05
1:A:843:HIS:O	1:A:845:PRO:HD3	1.59	1.02

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	268/295 (91%)	228 (85%)	30 (11%)	10 (4%)	5	20
1	D	263/295 (89%)	221 (84%)	35 (13%)	7 (3%)	8	30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	266/295 (90%)	223 (84%)	34 (13%)	9 (3%)	6	23
1	H	263/295 (89%)	226 (86%)	29 (11%)	8 (3%)	7	27
2	B	178/230 (77%)	135 (76%)	34 (19%)	9 (5%)	3	10
2	C	178/230 (77%)	138 (78%)	32 (18%)	8 (4%)	4	14
2	F	178/230 (77%)	136 (76%)	36 (20%)	6 (3%)	6	23
2	G	178/230 (77%)	138 (78%)	33 (18%)	7 (4%)	5	18
All	All	1772/2100 (84%)	1445 (82%)	263 (15%)	64 (4%)	5	22

5 of 64 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	845	PRO
1	D	849	ASN
1	E	845	PRO
1	H	874	SER
1	H	888	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	224/259 (86%)	196 (88%)	28 (12%)	7	19
1	D	222/259 (86%)	194 (87%)	28 (13%)	7	18
1	E	223/259 (86%)	188 (84%)	35 (16%)	4	11
1	H	222/259 (86%)	203 (91%)	19 (9%)	15	41
2	B	166/210 (79%)	153 (92%)	13 (8%)	18	46
2	C	166/210 (79%)	154 (93%)	12 (7%)	21	51
2	F	166/210 (79%)	155 (93%)	11 (7%)	24	57
2	G	166/210 (79%)	157 (95%)	9 (5%)	31	69
All	All	1555/1876 (83%)	1400 (90%)	155 (10%)	11	32

5 of 155 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	807	ASP
1	E	665	SER
1	H	694	TRP
1	D	816	GLU
1	D	855	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	896	HIS
1	E	896	HIS
1	H	817	HIS
1	E	753	ASN
2	B	313	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 ⓘ

4 ligands are 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
3	SAH	A	1	-	28,28,28	1.93	8 (28%)	40,40,40	2.60	11 (27%)
3	SAH	D	4	-	28,28,28	1.88	5 (17%)	40,40,40	2.52	11 (27%)
3	SAH	E	5	-	28,28,28	1.93	9 (32%)	40,40,40	2.23	11 (27%)
3	SAH	H	8	-	28,28,28	1.96	7 (25%)	40,40,40	2.78	12 (30%)

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	SAH	A	1	-	-	0/15/31/31	0/1/3/3
3	SAH	D	4	-	-	0/15/31/31	0/1/3/3
3	SAH	E	5	-	-	0/15/31/31	0/1/3/3
3	SAH	H	8	-	-	1/15/31/31	0/1/3/3

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	4	SAH	C1'-N9	-5.11	1.32	1.48
3	E	5	SAH	C1'-N9	-5.07	1.32	1.48
3	A	1	SAH	C1'-N9	-4.54	1.34	1.48
3	H	8	SAH	C1'-N9	-4.50	1.34	1.48
3	A	1	SAH	O4'-C1'	4.07	1.47	1.41

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1	SAH	O4'-C1'-C2'	-9.22	92.63	106.77
3	H	8	SAH	O4'-C1'-N9	8.43	116.28	108.44
3	E	5	SAH	N3-C2-N1	-8.37	121.71	128.71
3	A	1	SAH	N3-C2-N1	-7.65	122.31	128.71
3	H	8	SAH	N3-C2-N1	-7.55	122.39	128.71

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	H	8	SAH	C2'-C1'-N9-C8

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	272/295 (92%)	0.16	5 (1%) 65 74	30, 64, 106, 146	0
1	D	267/295 (90%)	0.05	4 (1%) 70 79	37, 75, 122, 157	0
1	E	270/295 (91%)	0.17	5 (1%) 64 72	27, 67, 111, 132	0
1	H	267/295 (90%)	0.52	35 (13%) 4 6	51, 101, 138, 170	0
2	B	186/230 (80%)	0.90	33 (17%) 2 3	55, 121, 187, 207	0
2	C	186/230 (80%)	1.05	40 (21%) 1 2	61, 118, 189, 206	0
2	F	186/230 (80%)	1.07	37 (19%) 2 2	66, 121, 189, 210	0
2	G	186/230 (80%)	2.53	84 (45%) 1 0	109, 163, 201, 210	0
All	All	1820/2100 (86%)	0.70	243 (13%) 4 5	27, 93, 181, 210	0

The worst 5 of 243 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	321	VAL	19.9
2	G	309	ILE	13.6
2	G	334	HIS	12.8
2	F	331	ARG	12.5
2	G	331	ARG	12.3

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
3	SAH	A	1	26/26	0.22	0.02	38,67,84,95	0
3	SAH	D	4	26/26	0.19	-0.28	40,59,71,73	0
3	SAH	H	8	26/26	0.17	-0.80	49,104,112,114	0
3	SAH	E	5	26/26	0.16	-0.82	31,56,66,69	0

6.5 Other polymers

There are no such residues in this entry.