



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 12, 2021 – 12:09 PM EDT

PDB ID : 7LNJ  
Title : CamA Adenine Methyltransferase Complexed to Cognate Substrate DNA  
Authors : Horton, J.R.; Cheng, X.; Zhou, J.  
Deposited on : 2021-02-07  
Resolution : 2.68 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](mailto: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  
Xtriage (Phenix) : 1.13  
EDS : 2.18  
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) : 2.18

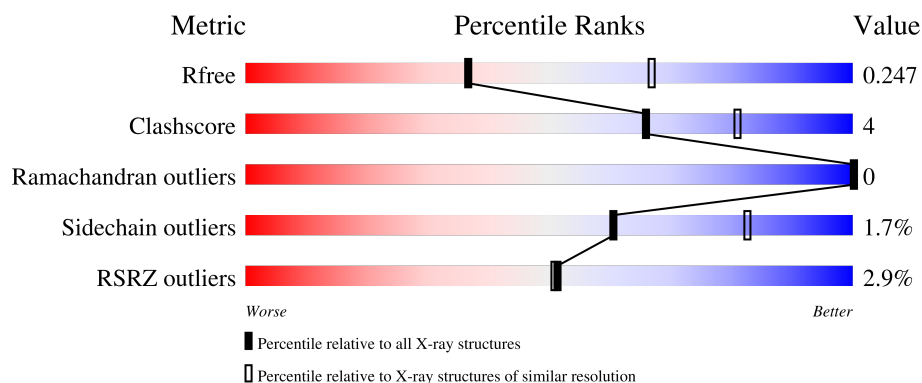
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.68 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	3863 (2.70-2.66)
Clashscore	141614	4210 (2.70-2.66)
Ramachandran outliers	138981	4141 (2.70-2.66)
Sidechain outliers	138945	4141 (2.70-2.66)
RSRZ outliers	127900	3780 (2.70-2.66)

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  $\geq 3$ , 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	578	<div> <div>81%</div> <div>12%</div> <div>6%</div> </div>
1	B	578	<div> <div>84%</div> <div>10%</div> <div>5%</div> </div>
1	C	578	<div> <div>5%</div> <div>81%</div> <div>12%</div> <div>7%</div> </div>
2	E	14	<div> <div>7%</div> <div>79%</div> <div>14%</div> <div>7%</div> </div>
2	G	14	<div> <div>71%</div> <div>29%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	I	14	
3	D	14	
3	F	14	
3	H	14	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 15057 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 Site-specific DNA-methyltransferase (adenine-specific).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	543	Total	C	N	O	S	0	1	0
			4465	2908	714	827	16			
1	B	550	Total	C	N	O	S	0	0	0
			4502	2932	720	834	16			
1	C	539	Total	C	N	O	S	0	0	0
			4280	2791	681	791	17			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	HIS	-	expression tag	UNP Q183J3
B	0	HIS	-	expression tag	UNP Q183J3
C	0	HIS	-	expression tag	UNP Q183J3

- Molecule 2 is a DNA chain called DNA Strand 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	13	Total	C	N	O	P	0	0	0
			269	129	45	82	13			
2	G	14	Total	C	N	O	P	0	0	0
			287	139	50	85	13			
2	I	14	Total	C	N	O	P	0	0	0
			287	139	50	85	13			

- Molecule 3 is a DNA chain called DNA Strand 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	14	Total	C	N	O	P	0	0	0
			281	136	53	79	13			
3	F	14	Total	C	N	O	P	0	0	0
			281	136	53	79	13			
3	H	14	Total	C	N	O	P	0	0	0
			281	136	53	79	13			

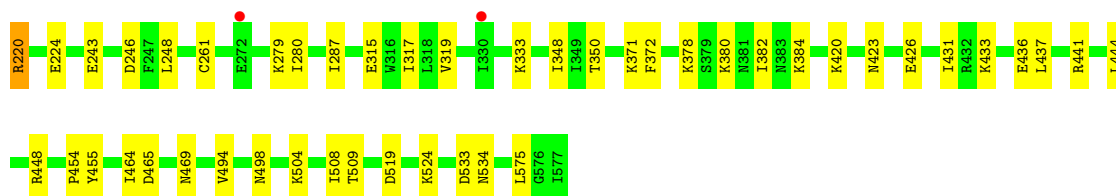
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	29	Total 29	O 29	0	0
4	B	40	Total 40	O 40	0	0
4	C	14	Total 14	O 14	0	0
4	E	3	Total 3	O 3	0	0
4	D	10	Total 10	O 10	0	0
4	F	11	Total 11	O 11	0	0
4	G	4	Total 4	O 4	0	0
4	H	11	Total 11	O 11	0	0
4	I	2	Total 2	O 2	0	0

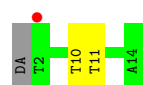
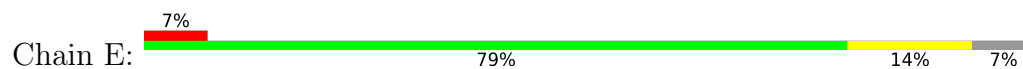


- Molecule 1: Site-specific DNA-methyltransferase (adenine-specific)

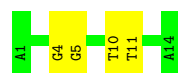




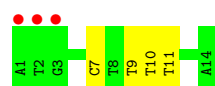
- Molecule 2: DNA Strand 2



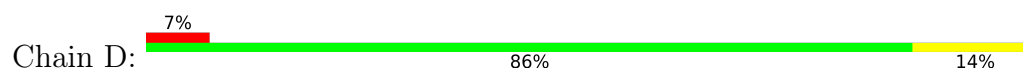
- Molecule 2: DNA Strand 2



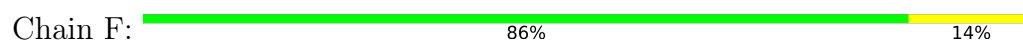
- Molecule 2: DNA Strand 2



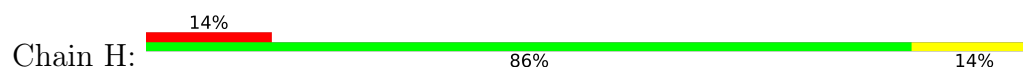
- Molecule 3: DNA Strand 1



- Molecule 3: DNA Strand 1



- Molecule 3: DNA Strand 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.25Å 161.36Å 230.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.35 – 2.68 44.35 – 2.68	Depositor EDS
% Data completeness (in resolution range)	87.3 (44.35-2.68) 87.3 (44.35-2.68)	Depositor EDS
$R_{merge}$	0.22	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.27 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, $R_{free}$	0.210 , 0.251 0.208 , 0.247	Depositor DCC
$R_{free}$ test set	3871 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.9	Xtriage
Anisotropy	0.520	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 39.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15057	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.47% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

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	0.24	0/4560	0.43	0/6144
1	B	0.24	0/4596	0.43	0/6197
1	C	0.24	0/4368	0.43	0/5904
2	E	0.52	0/300	1.01	0/462
2	G	0.54	0/321	1.00	0/495
2	I	0.52	0/321	0.98	0/495
3	D	0.46	0/315	0.84	0/483
3	F	0.45	0/315	0.83	0/483
3	H	0.45	0/315	0.82	0/483
All	All	0.29	0/15411	0.53	0/21146

There are no bond length outliers.

There are no bond angle outliers.

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	4465	0	4412	40	0
1	B	4502	0	4431	31	0
1	C	4280	0	4104	37	0
2	E	269	0	150	1	0
2	G	287	0	162	2	0
2	I	287	0	162	3	0
3	D	281	0	159	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	F	281	0	159	1	0
3	H	281	0	159	1	0
4	A	29	0	0	0	0
4	B	40	0	0	0	0
4	C	14	0	0	0	0
4	D	10	0	0	0	0
4	E	3	0	0	0	0
4	F	11	0	0	0	0
4	G	4	0	0	0	0
4	H	11	0	0	0	0
4	I	2	0	0	0	0
All	All	15057	0	13898	115	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 4.

The worst 5 of 115 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:349:ILE:HB	1:A:439:TRP:HB2	1.77	0.66
1:A:69:LEU:HD12	1:A:123:LEU:HD21	1.78	0.65
1:A:369:ASP:OD2	1:A:371:LYS:NZ	2.28	0.64
1:C:220:ARG:HB3	1:C:261:CYS:HB3	1.80	0.64
1:B:66:GLY:HA3	1:B:123:LEU:HD13	1.78	0.64

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	540/578 (93%)	526 (97%)	14 (3%)	0	100 100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	548/578 (95%)	535 (98%)	13 (2%)	0	100	100
1	C	535/578 (93%)	522 (98%)	13 (2%)	0	100	100
All	All	1623/1734 (94%)	1583 (98%)	40 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	484/548 (88%)	477 (99%)	7 (1%)	67	85
1	B	486/548 (89%)	477 (98%)	9 (2%)	57	80
1	C	443/548 (81%)	435 (98%)	8 (2%)	59	81
All	All	1413/1644 (86%)	1389 (98%)	24 (2%)	60	82

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	455	TYR
1	C	127	LEU
1	C	77	TYR
1	C	168	TYR
1	A	519	ASP

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

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	543/578 (93%)	-0.18	4 (0%) 87 88	30, 52, 84, 145	0
1	B	550/578 (95%)	-0.08	8 (1%) 73 74	32, 53, 87, 178	0
1	C	539/578 (93%)	0.22	31 (5%) 23 21	44, 70, 125, 158	0
2	E	13/14 (92%)	-0.32	1 (7%) 13 11	43, 45, 123, 159	0
2	G	14/14 (100%)	-0.60	0 100 100	41, 48, 66, 67	0
2	I	14/14 (100%)	0.55	3 (21%) 0 0	53, 57, 155, 159	0
3	D	14/14 (100%)	-0.15	1 (7%) 16 13	36, 45, 128, 153	0
3	F	14/14 (100%)	-0.17	0 100 100	36, 42, 77, 84	0
3	H	14/14 (100%)	0.36	2 (14%) 2 1	52, 62, 124, 158	0
All	All	1715/1818 (94%)	-0.02	50 (2%) 51 51	30, 58, 111, 178	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	H	14	DA	6.5
1	C	91	TYR	6.0
2	I	1	DA	5.6
1	C	87	LEU	4.8
1	C	95	TYR	4.8

### 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 monosaccharides 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.