



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 15, 2024 – 10:12 PM EST

PDB ID : 6U8V  
Title : Crystal structure of DNMT3B-DNMT3L in complex with CpGpT DNA  
Authors : Gao, L.; Zhang, Z.M.; Song, J.  
Deposited on : 2019-09-06  
Resolution : 3.00 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 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 : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.21  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

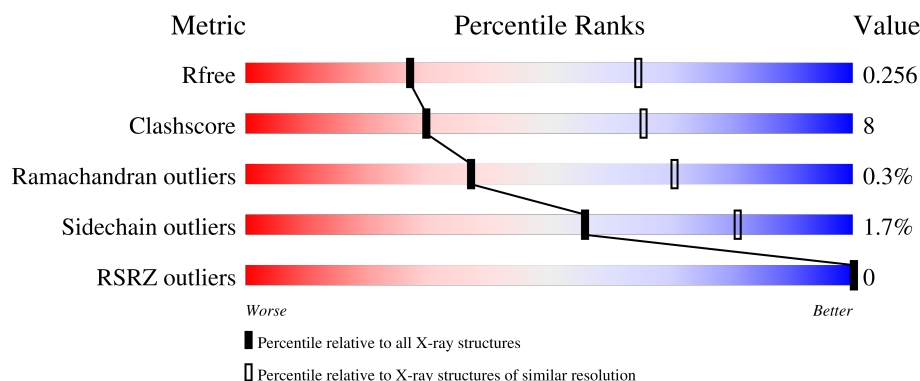
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*





The reported resolution of this entry is 3.00 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))
$R_{free}$	164625	2511 (3.00-3.00)
Clashscore	180529	2866 (3.00-3.00)
Ramachandran outliers	177936	2778 (3.00-3.00)
Sidechain outliers	177891	2781 (3.00-3.00)
RSRZ outliers	164620	2523 (3.00-3.00)

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	291	
1	D	291	
2	B	209	
2	C	209	
3	E	25	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	25	 A horizontal bar chart showing the quality of chain F. The bar is divided into three segments: a green segment representing 64%, a yellow segment representing 32%, and a small orange segment at the end. A single dot is visible at the far right end of the bar.

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8597 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 3B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	291	Total	C	N	O	S	0	1	0
			2303	1477	404	410	12			
1	D	291	Total	C	N	O	S	0	1	0
			2312	1485	404	411	12			

- Molecule 2 is a protein called DNA (cytosine-5)-methyltransferase 3-like.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	193	Total	C	N	O	S	0	0	0
			1443	944	244	252	3			
2	C	193	Total	C	N	O	S	0	0	0
			1443	941	245	253	4			

- Molecule 3 is a DNA chain called CpGpT DNA (25-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	25	Total	C	N	O	P	0	0	0
			510	245	93	148	24			
3	F	25	Total	C	N	O	P	0	0	0
			510	245	93	148	24			

- Molecule 4 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: C<sub>14</sub>H<sub>20</sub>N<sub>6</sub>O<sub>5</sub>S) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

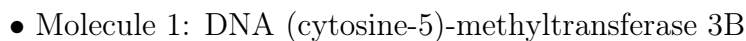
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Mg	0	0
			1	1		
5	D	1	Total	Mg	0	0
			1	1		

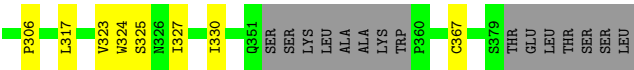
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	6	Total	O	0	0
			6	6		
6	D	15	Total	O	0	0
			15	15		
6	F	1	Total	O	0	0
			1	1		

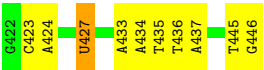


- Molecule 1: DNA (cytosine-5)-methyltransferase 3B

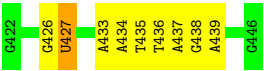




- Molecule 3: CpGpT DNA (25-MER)



- Molecule 3: CpGpT DNA (25-MER)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	193.76Å 193.76Å 49.76Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.55 – 3.00 48.55 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.3 (48.55-3.00) 99.6 (48.55-3.00)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.92 (at 3.01Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.218 , 0.249 0.224 , 0.256	Depositor DCC
$R_{free}$ test set	39815 reflections (4.73%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	68.5	Xtriage
Anisotropy	0.457	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 48.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.030 for -h,-k,l 0.478 for h,-h-k,-l 0.027 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8597	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PYO, MG, 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	0.27	0/2357	0.44	0/3191
1	D	0.27	0/2367	0.45	0/3204
2	B	0.25	0/1491	0.43	0/2046
2	C	0.27	0/1491	0.47	0/2048
3	E	0.49	0/550	0.88	0/846
3	F	0.48	0/550	0.90	0/846
All	All	0.30	0/8806	0.53	0/12181

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
1	D	0	1
2	C	0	2
All	All	0	4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	625	VAL	Peptide
2	C	210	GLU	Peptide
2	C	317	LEU	Peptide
1	D	625	VAL	Peptide

## 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	2303	0	2277	45	0
1	D	2312	0	2288	38	0
2	B	1443	0	1280	22	0
2	C	1443	0	1276	28	0
3	E	510	0	278	7	0
3	F	510	0	279	10	0
4	A	26	0	19	0	0
4	D	26	0	19	0	0
5	A	1	0	0	0	0
5	D	1	0	0	0	0
6	A	6	0	0	0	0
6	D	15	0	0	0	0
6	F	1	0	0	0	0
All	All	8597	0	7716	136	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 8.

The worst 5 of 136 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)
2:C:224:VAL:HG11	2:C:266:LEU:HD11	1.58	0.83
2:B:240:LEU:HD13	2:B:371:LEU:HD11	1.68	0.75
1:D:782:LYS:HB3	1:D:782:LYS:NZ	2.01	0.74
1:D:761:GLU:OE1	1:D:807:ARG:NH1	2.22	0.73
1:A:754:LEU:O	1:A:770:LYS:NZ	2.22	0.71

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	290/291 (100%)	283 (98%)	6 (2%)	1 (0%)	37	70
1	D	290/291 (100%)	282 (97%)	6 (2%)	2 (1%)	19	54
2	B	189/209 (90%)	183 (97%)	6 (3%)	0	100	100
2	C	189/209 (90%)	177 (94%)	12 (6%)	0	100	100
All	All	958/1000 (96%)	925 (97%)	30 (3%)	3 (0%)	37	70

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	626	ASN
1	D	626	ASN
1	D	627	ASP

### 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	242/252 (96%)	237 (98%)	5 (2%)	48	77
1	D	243/252 (96%)	241 (99%)	2 (1%)	79	90
2	B	133/191 (70%)	131 (98%)	2 (2%)	60	83
2	C	134/191 (70%)	130 (97%)	4 (3%)	36	69
All	All	752/886 (85%)	739 (98%)	13 (2%)	56	81

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	194	SER
2	C	224	VAL
1	D	788	LEU
2	C	367	CYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	782	LYS

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 ⓘ

2 non-standard protein/DNA/RNA residues 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	PYO	E	427	3	16,20,21	4.74	13 (81%)	21,28,31	1.34	3 (14%)
3	PYO	F	427	3	16,20,21	2.61	5 (31%)	21,28,31	0.91	1 (4%)

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	PYO	E	427	3	-	4/7/25/26	0/2/2/2
3	PYO	F	427	3	-	4/7/25/26	0/2/2/2

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	427	PYO	C2'-C3'	-10.42	1.25	1.53
3	E	427	PYO	C6-C5	7.07	1.50	1.35

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	427	PYO	C6-C5	7.05	1.50	1.35
3	E	427	PYO	O4'-C4'	-6.82	1.29	1.45
3	E	427	PYO	C3'-C4'	5.34	1.66	1.53

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	427	PYO	C5-C4-N3	-3.71	119.50	124.25
3	F	427	PYO	C5-C4-N3	-3.35	119.96	124.25
3	E	427	PYO	C4'-O4'-C1'	-3.06	102.70	109.47
3	E	427	PYO	C3'-C2'-C1'	2.31	105.84	101.46

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	427	PYO	C2'-C1'-N1-C6
3	F	427	PYO	C2'-C1'-N1-C6
3	E	427	PYO	C2'-C1'-N1-C2
3	F	427	PYO	C2'-C1'-N1-C2
3	F	427	PYO	O4'-C1'-N1-C2

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	427	PYO	2	0
3	F	427	PYO	4	0

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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
4	SAH	D	901	-	23,28,28	1.28	3 (13%)	22,40,40	1.87	4 (18%)
4	SAH	A	901	-	23,28,28	1.28	3 (13%)	22,40,40	1.94	4 (18%)

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
4	SAH	D	901	-	-	5/11/31/31	0/3/3/3
4	SAH	A	901	-	-	5/11/31/31	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	901	SAH	C2-N3	4.24	1.38	1.32
4	D	901	SAH	C2-N3	4.24	1.38	1.32
4	D	901	SAH	C2-N1	2.65	1.38	1.33
4	A	901	SAH	C2-N1	2.60	1.38	1.33
4	D	901	SAH	OXT-C	-2.12	1.23	1.30

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	901	SAH	N3-C2-N1	-6.33	120.07	128.67
4	D	901	SAH	N3-C2-N1	-6.29	120.14	128.67
4	A	901	SAH	C5'-SD-CG	-3.69	91.31	102.26
4	D	901	SAH	C5'-SD-CG	-3.25	92.60	102.26
4	A	901	SAH	OXT-C-O	-2.83	117.67	124.08

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	901	SAH	N-CA-CB-CG
4	D	901	SAH	N-CA-CB-CG

*Continued on next page...*

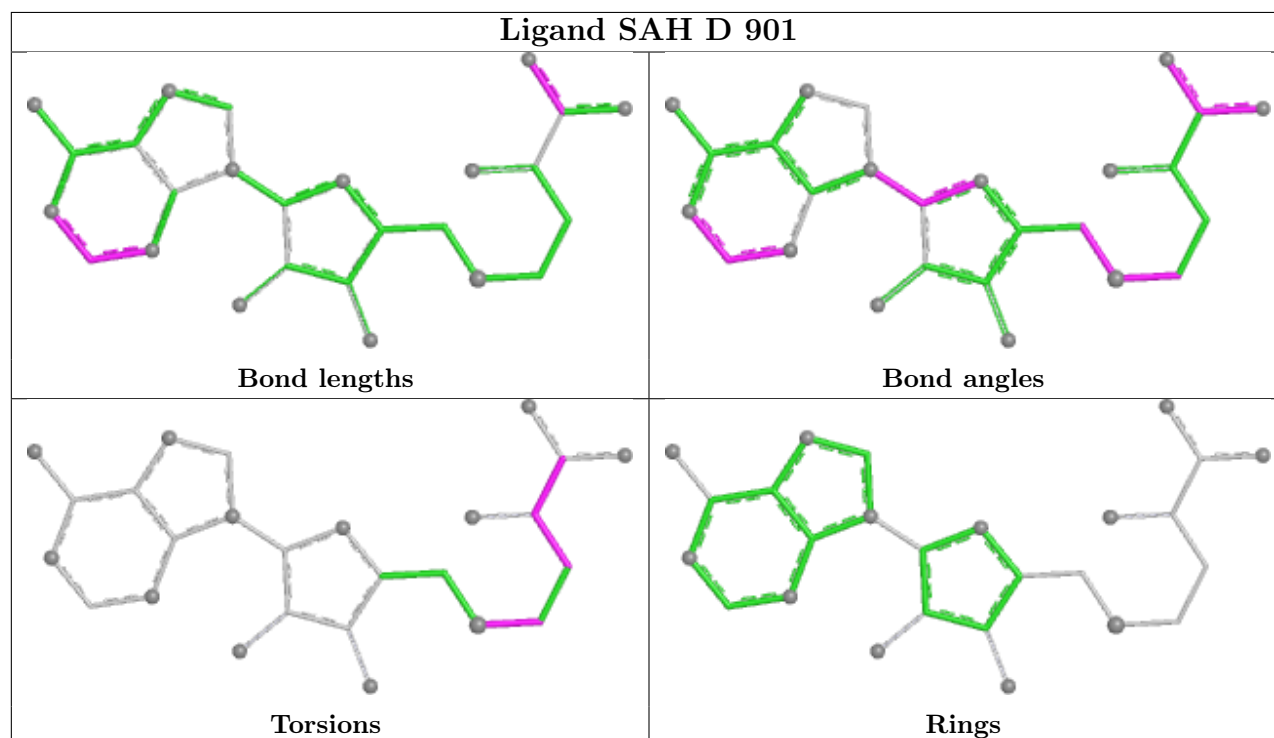
*Continued from previous page...*

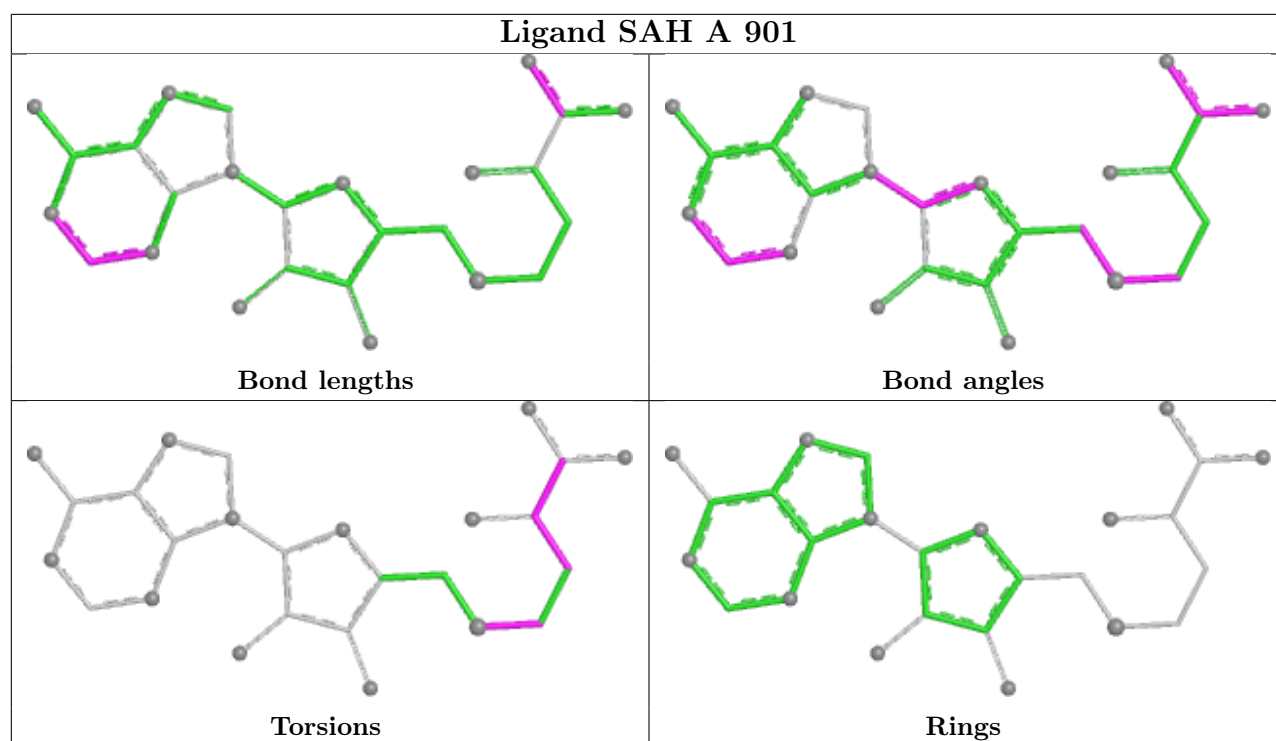
Mol	Chain	Res	Type	Atoms
4	D	901	SAH	C-CA-CB-CG
4	A	901	SAH	CB-CG-SD-C5'
4	A	901	SAH	O-C-CA-CB

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, 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	291/291 (100%)	-1.54	0 100 100	26, 59, 95, 138	1 (0%)
1	D	291/291 (100%)	-1.54	0 100 100	26, 59, 98, 151	1 (0%)
2	B	193/209 (92%)	-1.15	0 100 100	67, 110, 180, 187	0
2	C	193/209 (92%)	-1.12	0 100 100	70, 111, 175, 219	0
3	E	24/25 (96%)	-1.05	0 100 100	88, 135, 162, 169	0
3	F	24/25 (96%)	-1.21	0 100 100	90, 128, 155, 168	0
All	All	1016/1050 (96%)	-1.37	0 100 100	26, 76, 162, 219	2 (0%)

There are no RSRZ outliers to report.

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	PYO	E	427	19/20	0.99	0.04	52,67,81,84	0
3	PYO	F	427	19/20	0.99	0.03	44,71,79,87	0

### 6.3 Carbohydrates [i](#)

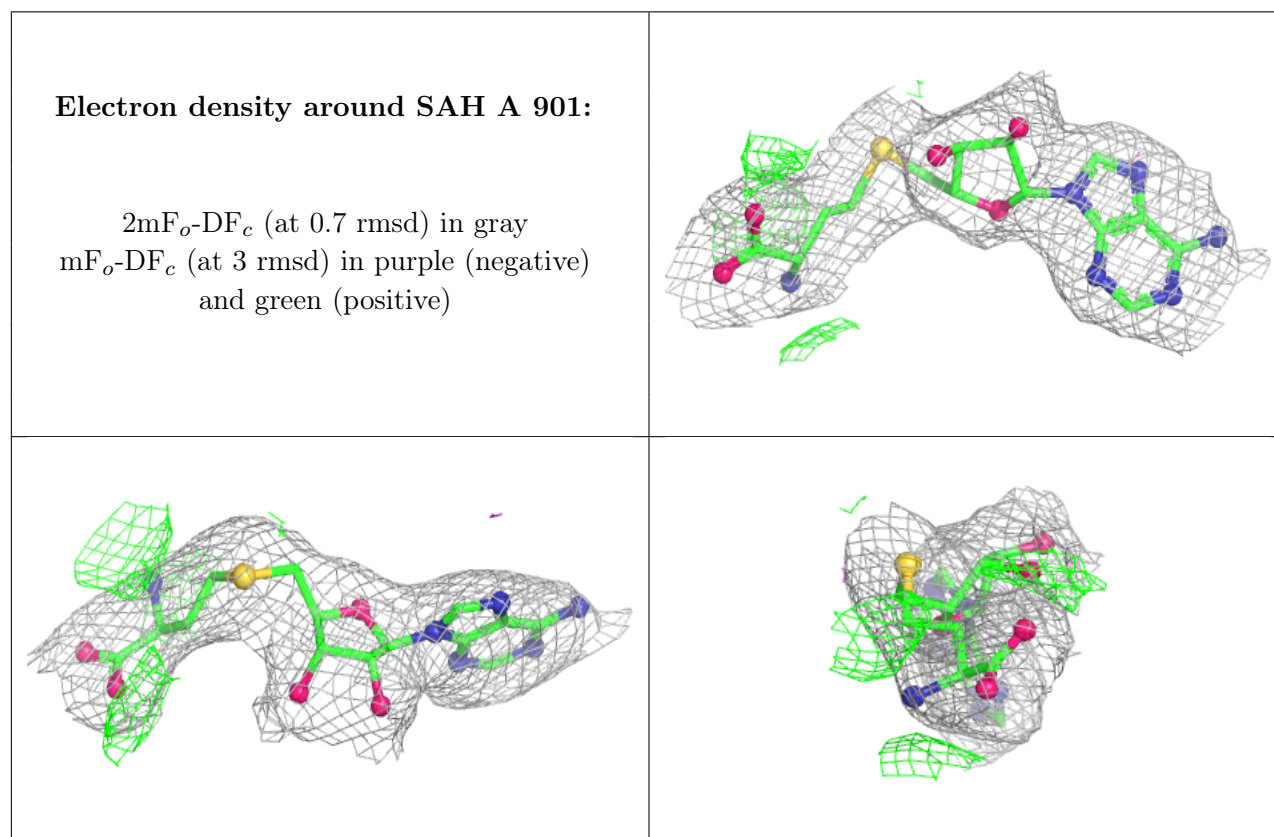
There are no monosaccharides in this entry.

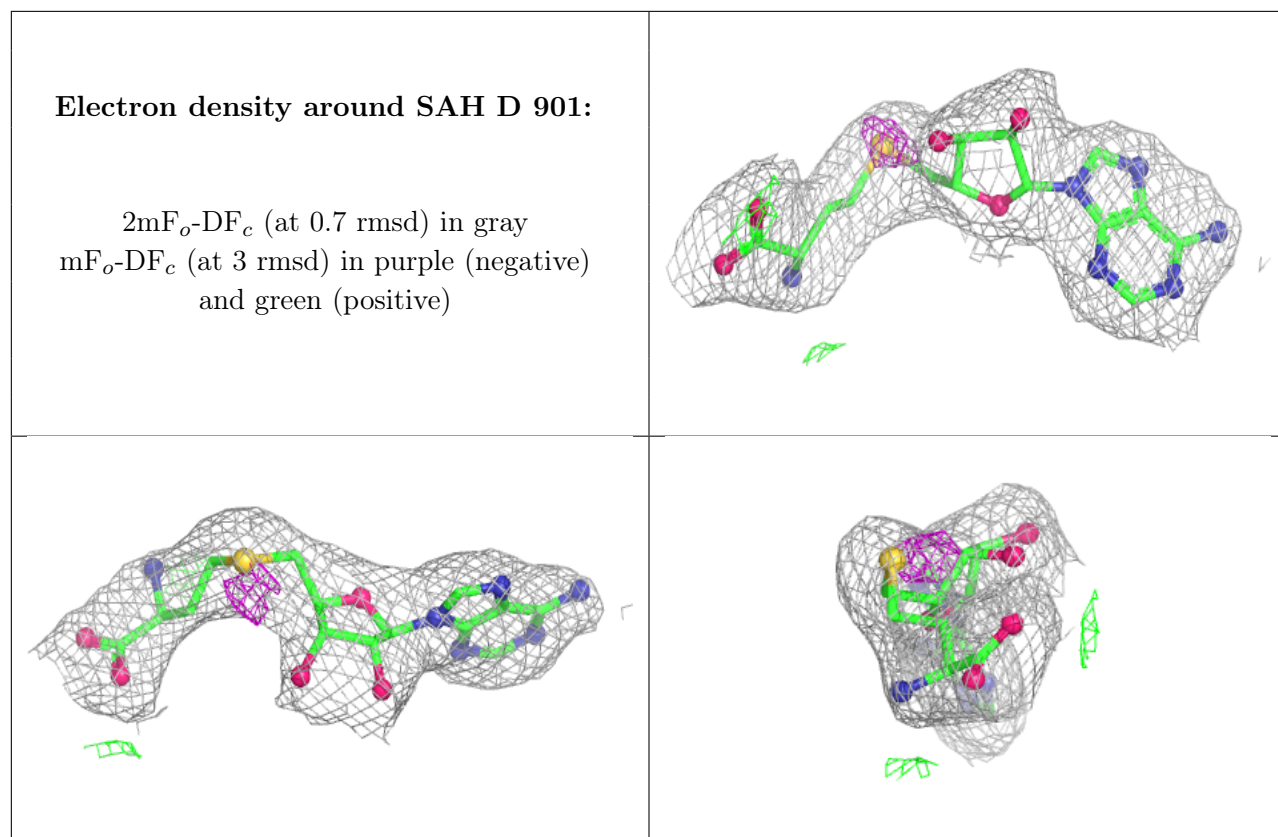
## 6.4 Ligands [i](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
4	SAH	A	901	26/26	0.99	0.03	25,40,48,56	0
4	SAH	D	901	26/26	0.99	0.03	24,42,49,59	0
5	MG	A	902	1/1	0.99	0.09	65,65,65,65	0
5	MG	D	902	1/1	0.99	0.07	70,70,70,70	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.