



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 31, 2023 – 04:24 PM JST

PDB ID : 5GUT  
Title : The crystal structure of mouse DNMT1 (731-1602) mutant - N1248A  
Authors : Chen, S.J.; Ye, F.  
Deposited on : 2016-08-31  
Resolution : 2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

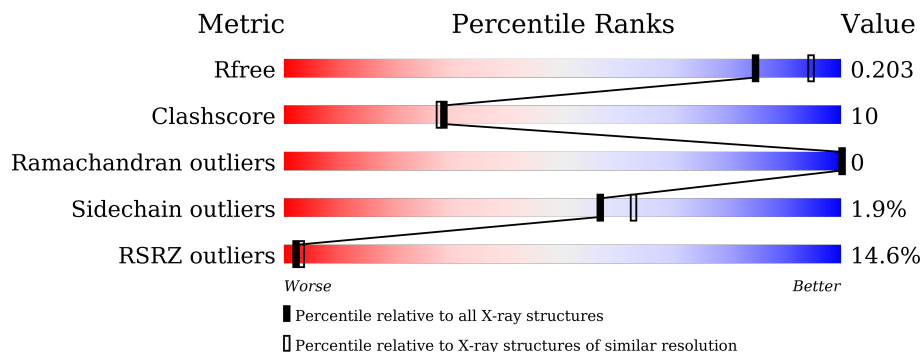
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 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	872	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 6962 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.

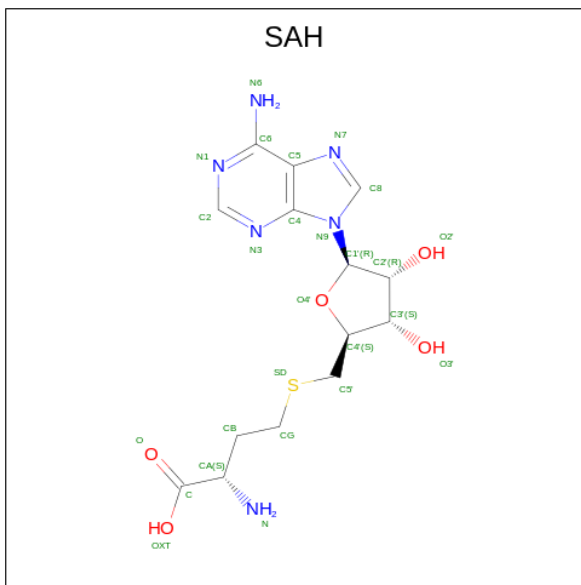
- Molecule 1 is a protein called DNA (cytosine-5)-methyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	794	6405	4071	1124	1168	42	0	7	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1248	ALA	ASN	engineered mutation	UNP P13864

- Molecule 2 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).



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

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		

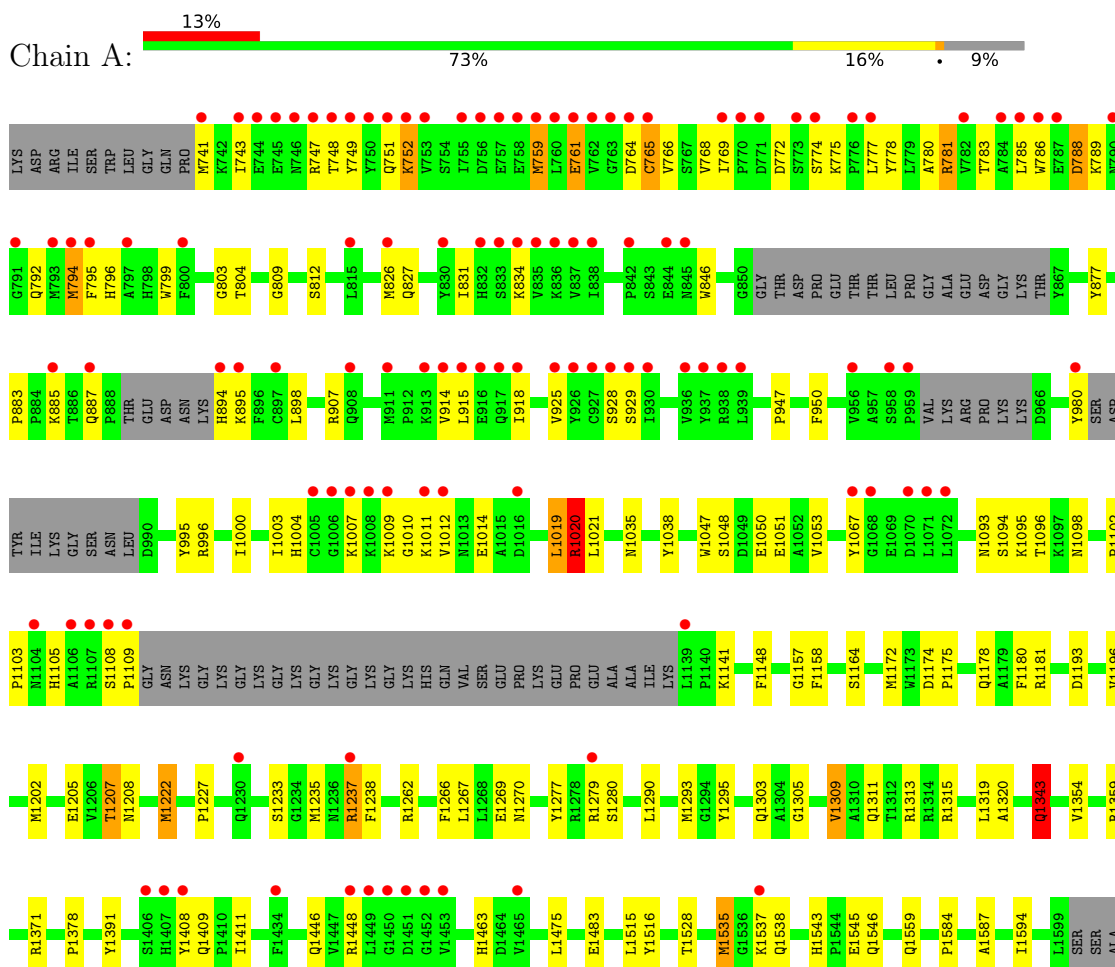
- Molecule 5 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	A	479	Total 479	O 479	0	0

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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)-methyltransferase 1



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.42Å 78.55Å 164.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.36 – 2.10 44.36 – 2.10	Depositor EDS
% Data completeness (in resolution range)	96.6 (44.36-2.10) 96.6 (44.36-2.10)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.36 (at 2.10Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.200 , 0.231 0.202 , 0.203	Depositor DCC
$R_{free}$ test set	2839 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.7	Xtrriage
Anisotropy	0.042	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 66.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.022 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6962	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.23% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, SAH, 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	0.58	0/6593	0.88	19/8925 (0.2%)

There are no bond length outliers.

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	781	ARG	CG-CD-NE	-9.32	92.22	111.80
1	A	1408	TYR	CB-CG-CD1	-7.31	116.61	121.00
1	A	1019	LEU	CA-CB-CG	6.88	131.12	115.30
1	A	794	MET	CB-CG-SD	-6.85	91.85	112.40
1	A	1343	GLN	CA-CB-CG	-6.76	98.52	113.40
1	A	1408	TYR	CB-CG-CD2	6.19	124.72	121.00
1	A	788	ASP	CB-CG-OD1	-6.11	112.80	118.30
1	A	1172	MET	CG-SD-CE	6.08	109.93	100.20
1	A	1020	ARG	CG-CD-NE	5.98	124.36	111.80
1	A	1180	PHE	CB-CG-CD2	-5.72	116.80	120.80
1	A	1019	LEU	CB-CG-CD1	-5.63	101.42	111.00
1	A	761	GLU	CA-CB-CG	5.39	125.26	113.40
1	A	1235	MET	C-N-CA	-5.35	108.33	121.70
1	A	1475	LEU	CB-CG-CD2	5.34	120.07	111.00
1	A	1309	VAL	CG1-CB-CG2	5.32	119.40	110.90
1	A	1535	MET	C-N-CA	-5.23	111.31	122.30
1	A	907	ARG	CG-CD-NE	5.23	122.78	111.80
1	A	1180	PHE	CB-CG-CD1	5.20	124.44	120.80
1	A	765	CYS	CA-CB-SG	-5.01	104.98	114.00

There are no chirality outliers.

There are no planarity outliers.



## 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	6405	0	6262	124	1
2	A	26	0	19	0	0
3	A	2	0	0	0	0
4	A	50	0	0	2	1
5	A	479	0	0	13	1
All	All	6962	0	6281	124	2

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 10.

All (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:1537:LYS:NZ	5:A:1803:HOH:O	2.00	0.95
1:A:1303:GLN:HE22	1:A:1313:ARG:H	1.11	0.94
1:A:1371:ARG:NH1	5:A:1804:HOH:O	2.01	0.93
1:A:1237:ARG:HH21	1:A:1279:ARG:HB2	1.39	0.88
1:A:1178:GLN:OE1	5:A:1802:HOH:O	1.91	0.88
1:A:741:MET:HG3	1:A:751:GLN:HG2	1.55	0.87
1:A:1303:GLN:HE21	1:A:1305:GLY:H	1.21	0.84
1:A:765:CYS:SG	1:A:781:ARG:NH1	2.52	0.82
1:A:1237:ARG:HE	1:A:1279:ARG:HD3	1.46	0.80
1:A:789:LYS:O	5:A:1805:HOH:O	2.04	0.75
1:A:1237:ARG:HG2	1:A:1279:ARG:HH11	1.52	0.74
1:A:1311:GLN:HE22	1:A:1313:ARG:HH11	1.36	0.74
1:A:772:ASP:OD1	1:A:774:SER:OG	2.10	0.67
1:A:1067:TYR:CE1	1:A:1102:PRO:HG2	2.30	0.66
1:A:1237:ARG:NH2	5:A:1806:HOH:O	2.11	0.66
1:A:1178:GLN:NE2	5:A:1809:HOH:O	2.28	0.65
1:A:752:LYS:HG3	1:A:759:MET:CE	2.26	0.65
1:A:1303:GLN:NE2	1:A:1313:ARG:H	1.91	0.65
1:A:1096:THR:O	1:A:1098:ASN:ND2	2.29	0.65
1:A:1543:HIS:HD2	1:A:1545:GLU:H	1.45	0.64
1:A:1237:ARG:CG	1:A:1279:ARG:HH11	2.10	0.64
1:A:1311:GLN:NE2	1:A:1528:THR:H	1.95	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1000:ILE:HG23	1:A:1019:LEU:HD12	1.81	0.62
1:A:780:ALA:HB2	1:A:799:TRP:CE3	2.34	0.62
1:A:914:VAL:HG23	1:A:1014:GLU:HG2	1.81	0.62
1:A:747:ARG:HD3	1:A:786:TRP:NE1	2.14	0.61
1:A:1290:LEU:HD23	1:A:1293:MET:HE3	1.83	0.60
1:A:752:LYS:HG3	1:A:759:MET:SD	2.42	0.59
1:A:1343:GLN:NE2	5:A:1815:HOH:O	2.36	0.59
1:A:1181:ARG:HD2	5:A:1802:HOH:O	2.02	0.59
1:A:1050:GLU:OE1	1:A:1094:SER:HB2	2.03	0.58
1:A:766:VAL:HG13	1:A:831:ILE:HG23	1.86	0.58
1:A:1543:HIS:HE1	5:A:1882:HOH:O	1.86	0.58
1:A:1463:HIS:O	1:A:1483:GLU:HG2	2.04	0.56
1:A:1543:HIS:CD2	1:A:1545:GLU:H	2.23	0.56
1:A:747:ARG:HD3	1:A:786:TRP:CD1	2.41	0.56
1:A:741:MET:N	1:A:749:TYR:O	2.38	0.56
1:A:1207:THR:HG23	1:A:1208:ASN:O	2.05	0.56
1:A:788:ASP:CG	1:A:792:GLN:HB2	2.27	0.55
1:A:1537:LYS:HD2	1:A:1537:LYS:O	2.05	0.55
1:A:761:GLU:HG2	1:A:764:ASP:CG	2.27	0.55
1:A:1267:LEU:HD23	1:A:1319:LEU:HD23	1.89	0.55
1:A:995:TYR:OH	1:A:1359:ARG:HG2	2.07	0.55
1:A:788:ASP:OD1	1:A:792:GLN:HB2	2.05	0.55
1:A:1157:GLY:HA3	1:A:1587:ALA:HB3	1.87	0.55
1:A:1222:MET:CE	1:A:1594:ILE:HD13	2.37	0.55
1:A:1409:GLN:NE2	1:A:1411:ILE:O	2.40	0.54
1:A:1193:ASP:HB3	1:A:1196:VAL:HG13	1.89	0.54
1:A:1311:GLN:NE2	1:A:1313:ARG:HH11	2.05	0.54
1:A:1093:ASN:OD1	1:A:1095:LYS:HB2	2.08	0.54
1:A:1267:LEU:HD23	1:A:1319:LEU:CD2	2.38	0.54
1:A:918:ILE:HD11	1:A:928:SER:HB3	1.90	0.53
1:A:1543:HIS:CD2	1:A:1546:GLN:H	2.26	0.53
1:A:749:TYR:CD2	1:A:786:TRP:HB3	2.44	0.52
1:A:1009:LYS:HA	1:A:1009:LYS:HE2	1.91	0.52
1:A:1222:MET:HE3	1:A:1594:ILE:HD13	1.93	0.51
1:A:761:GLU:HG2	1:A:764:ASP:OD2	2.10	0.51
1:A:1543:HIS:HD2	1:A:1546:GLN:H	1.58	0.51
1:A:795:PHE:CE1	1:A:826:MET:HB3	2.46	0.51
1:A:1378:PRO:HG3	1:A:1391:TYR:O	2.12	0.50
1:A:1141:LYS:HE3	5:A:1993:HOH:O	2.10	0.50
1:A:1205:GLU:O	5:A:1807:HOH:O	2.19	0.50
1:A:914:VAL:HG21	1:A:1012:VAL:HG11	1.93	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1035:ASN:HA	1:A:1038:TYR:CZ	2.48	0.49
1:A:1516:TYR:CE2	1:A:1535:MET:HG3	2.47	0.49
1:A:747:ARG:CD	1:A:786:TRP:CD1	2.96	0.49
1:A:1103:PRO:HB2	1:A:1105:HIS:CE1	2.48	0.48
1:A:877:TYR:HA	1:A:1354:VAL:O	2.14	0.48
1:A:1266:PHE:HB3	1:A:1320:ALA:HB3	1.96	0.48
1:A:788:ASP:OD1	1:A:792:GLN:O	2.31	0.48
1:A:883:PRO:HD2	1:A:885:LYS:NZ	2.29	0.48
1:A:1537:LYS:HD2	1:A:1537:LYS:C	2.34	0.47
1:A:1007:LYS:HD2	1:A:1010:GLY:HA2	1.96	0.47
1:A:775:LYS:HD3	1:A:778:TYR:HE1	1.79	0.47
1:A:846:TRP:HH2	1:A:1295:TYR:CZ	2.33	0.47
1:A:1233:SER:HA	1:A:1277:TYR:CD2	2.50	0.47
1:A:1020:ARG:HH21	1:A:1053:VAL:HG21	1.80	0.47
1:A:761:GLU:O	1:A:764:ASP:HB2	2.15	0.47
1:A:1237:ARG:NE	1:A:1279:ARG:HD3	2.24	0.46
1:A:887:GLN:H	1:A:887:GLN:HG2	1.60	0.46
1:A:1515:LEU:HD11	1:A:1537:LYS:O	2.15	0.46
1:A:1516:TYR:HE2	1:A:1535:MET:HG3	1.81	0.46
1:A:775:LYS:HD3	1:A:778:TYR:CE1	2.51	0.46
1:A:768:VAL:HG11	1:A:826:MET:HE1	1.98	0.46
1:A:1003:ILE:HG12	1:A:1019:LEU:HD13	1.98	0.45
1:A:1237:ARG:CD	1:A:1279:ARG:NH1	2.79	0.45
1:A:1303:GLN:NE2	1:A:1305:GLY:H	2.02	0.45
1:A:950:PHE:HA	4:A:1712:SO4:O3	2.17	0.45
1:A:894:HIS:CE1	1:A:895:LYS:HE3	2.51	0.45
1:A:769:ILE:HD13	1:A:777:LEU:CD2	2.47	0.45
1:A:915:LEU:HB2	1:A:929:SER:OG	2.17	0.44
1:A:1047:TRP:CG	1:A:1048:SER:N	2.85	0.44
1:A:1020:ARG:HH21	1:A:1053:VAL:CG2	2.30	0.44
1:A:1108:SER:HB3	1:A:1109:PRO:HD2	2.00	0.44
1:A:1446:GLN:NE2	4:A:1706:SO4:O3	2.49	0.43
1:A:747:ARG:CZ	1:A:749:TYR:OH	2.67	0.43
1:A:743:ILE:HG22	1:A:748:THR:HG23	2.01	0.43
1:A:1011:LYS:HE3	1:A:1012:VAL:O	2.19	0.43
1:A:1237:ARG:CG	1:A:1279:ARG:NH1	2.81	0.43
1:A:751:GLN:O	1:A:785:LEU:HD12	2.19	0.42
1:A:1262:ARG:HA	1:A:1262:ARG:HD3	1.88	0.42
1:A:898:LEU:HD23	1:A:898:LEU:HA	1.87	0.42
1:A:1158:PHE:HB3	1:A:1164:SER:OG	2.19	0.42
1:A:1269:GLU:OE2	1:A:1315:ARG:NE	2.49	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1021:LEU:O	1:A:1051:GLU:HA	2.20	0.42
1:A:783:THR:OG1	1:A:796:HIS:HB3	2.20	0.41
1:A:794:MET:SD	1:A:827:GLN:HG2	2.59	0.41
1:A:1311:GLN:HE22	1:A:1313:ARG:NH1	2.13	0.41
1:A:1303:GLN:HE22	1:A:1313:ARG:N	1.95	0.41
1:A:794:MET:HE3	1:A:794:MET:HB3	1.84	0.41
1:A:1012:VAL:HG12	1:A:1014:GLU:HG3	2.02	0.41
1:A:947:PRO:HA	1:A:996:ARG:HG2	2.03	0.41
1:A:980:TYR:HD1	1:A:980:TYR:HA	1.63	0.41
1:A:1280:SER:HA	5:A:1950:HOH:O	2.19	0.41
1:A:1222:MET:HE1	1:A:1594:ILE:HD13	2.02	0.41
1:A:1237:ARG:HH21	1:A:1279:ARG:CB	2.19	0.41
1:A:1174:ASP:HB3	1:A:1175:PRO:HD3	2.02	0.41
1:A:1148:PHE:CE2	1:A:1227:PRO:HB3	2.56	0.40
1:A:768:VAL:HG11	1:A:826:MET:CE	2.52	0.40
1:A:803:GLY:HA3	1:A:812:SER:OG	2.22	0.40
1:A:1559:GLN:O	1:A:1584:PRO:HD2	2.21	0.40
1:A:804:THR:HA	1:A:809:GLY:O	2.22	0.40
1:A:1538:GLN:NE2	5:A:1819:HOH:O	2.54	0.40
1:A:925:VAL:O	1:A:1004:HIS:HA	2.21	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:2162:HOH:O	5:A:2196:HOH:O[4_455]	1.76	0.44
1:A:1095:LYS:NZ	4:A:1713:SO4:O2[3_554]	2.06	0.14

## 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	789/872 (90%)	756 (96%)	33 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	701/757 (93%)	688 (98%)	13 (2%)	57	63

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	752	LYS
1	A	759	MET
1	A	834	LYS
1	A	1020	ARG
1	A	1202	MET
1	A	1207	THR
1	A	1222	MET
1	A	1237	ARG
1	A	1238	PHE
1	A	1270	ASN
1	A	1309	VAL
1	A	1343	GLN
1	A	1448	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	894	HIS
1	A	1098	ASN
1	A	1160	GLN
1	A	1195	ASN
1	A	1303	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1311	GLN
1	A	1381	GLN
1	A	1543	HIS
1	A	1559	GLN

### 5.3.3 RNA [i](#)

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](#)

Of 13 ligands modelled in this entry, 2 are monoatomic - leaving 11 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	SO4	A	1705	-	4,4,4	0.14	0	6,6,6	0.11	0
2	SAH	A	1701	-	24,28,28	1.28	3 (12%)	25,40,40	1.76	3 (12%)
4	SO4	A	1709	-	4,4,4	0.14	0	6,6,6	0.36	0
4	SO4	A	1708	-	4,4,4	0.17	0	6,6,6	0.28	0
4	SO4	A	1713	-	4,4,4	0.11	0	6,6,6	0.11	0
4	SO4	A	1711	-	4,4,4	0.18	0	6,6,6	0.12	0
4	SO4	A	1704	-	4,4,4	0.09	0	6,6,6	0.28	0
4	SO4	A	1706	-	4,4,4	0.13	0	6,6,6	0.17	0
4	SO4	A	1712	-	4,4,4	0.19	0	6,6,6	0.30	0
4	SO4	A	1707	-	4,4,4	0.14	0	6,6,6	0.22	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	SO4	A	1710	1	4,4,4	0.13	0	6,6,6	0.18	0

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
2	SAH	A	1701	-	-	0/11/31/31	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1701	SAH	C2-N3	4.06	1.38	1.32
2	A	1701	SAH	C2-N1	2.73	1.39	1.33
2	A	1701	SAH	OXT-C	-2.33	1.22	1.30

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1701	SAH	N3-C2-N1	-5.60	119.93	128.68
2	A	1701	SAH	O4'-C1'-C2'	-3.87	101.27	106.93
2	A	1701	SAH	C4-C5-N7	-2.54	106.75	109.40

There are no chirality outliers.

There are no torsion outliers.

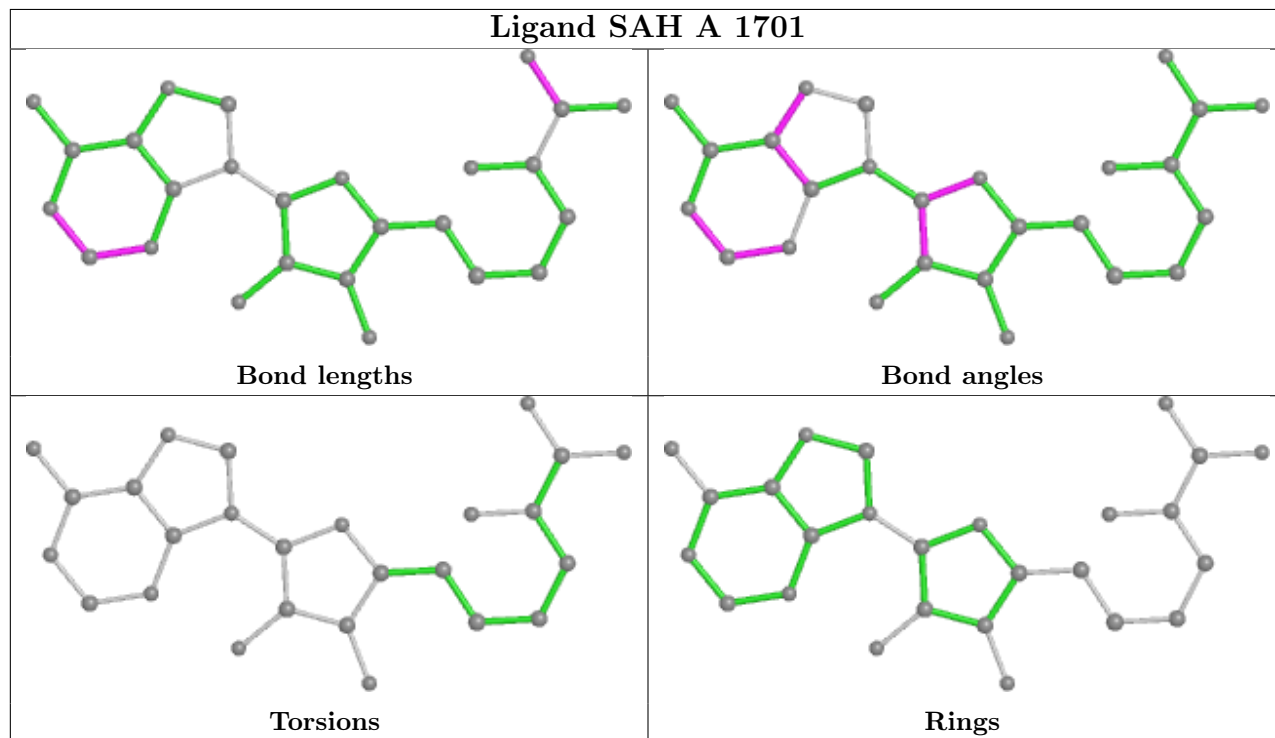
There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1713	SO4	0	1
4	A	1706	SO4	1	0
4	A	1712	SO4	1	0

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

### 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, 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	794/872 (91%)	0.69	116 (14%) <b>2</b> <b>3</b>	13, 32, 85, 130	0

All (116) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	750	TYR	7.9
1	A	785	LEU	7.5
1	A	741	MET	7.4
1	A	760	LEU	7.4
1	A	773	SER	7.2
1	A	1072	LEU	7.2
1	A	1451	ASP	6.8
1	A	761	GLU	6.5
1	A	755	ILE	6.4
1	A	745	GLU	6.0
1	A	1108	SER	5.9
1	A	748	THR	5.6
1	A	1406	SER	5.5
1	A	1106	ALA	5.2
1	A	749	TYR	5.2
1	A	918	ILE	5.1
1	A	1453	VAL	5.0
1	A	895	LYS	4.8
1	A	757	GLU	4.8
1	A	833	SER	4.8
1	A	1450	GLY	4.7
1	A	838	ILE	4.6
1	A	980	TYR	4.6
1	A	746	ASN	4.6
1	A	793	MET	4.5
1	A	763	GLY	4.5
1	A	915	LEU	4.4

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	786	TRP	4.4
1	A	753	VAL	4.4
1	A	1011	LYS	4.4
1	A	769	ILE	4.3
1	A	914	VAL	4.3
1	A	774	SER	4.2
1	A	936	VAL	4.2
1	A	1009	LYS	4.2
1	A	1109	PRO	4.1
1	A	752	LYS	4.1
1	A	759	MET	4.1
1	A	845	ASN	4.1
1	A	1452	GLY	4.1
1	A	830	TYR	4.1
1	A	1067	TYR	3.9
1	A	1005	CYS	3.9
1	A	1407	HIS	3.9
1	A	958	SER	3.8
1	A	758	GLU	3.8
1	A	887	GLN	3.8
1	A	1408	TYR	3.6
1	A	784	ALA	3.6
1	A	744	GLU	3.6
1	A	837	VAL	3.6
1	A	770	PRO	3.6
1	A	959	PRO	3.6
1	A	938	ARG	3.4
1	A	834	LYS	3.4
1	A	743	ILE	3.4
1	A	751	GLN	3.2
1	A	1008	LYS	3.2
1	A	776	PRO	3.2
1	A	1537	LYS	3.2
1	A	1071	LEU	3.1
1	A	1237	ARG	3.1
1	A	795	PHE	3.1
1	A	762	VAL	3.1
1	A	926	TYR	3.0
1	A	1012	VAL	3.0
1	A	930	ILE	3.0
1	A	747	ARG	3.0
1	A	928	SER	3.0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	927	CYS	2.9
1	A	885	LYS	2.9
1	A	835	VAL	2.9
1	A	956	VAL	2.9
1	A	1104	ASN	2.8
1	A	1139	LEU	2.8
1	A	790	ASN	2.7
1	A	939	LEU	2.7
1	A	1068	GLY	2.7
1	A	916	GLU	2.7
1	A	765	CYS	2.6
1	A	908	GLN	2.6
1	A	911	MET	2.6
1	A	782	VAL	2.6
1	A	836	LYS	2.6
1	A	897	CYS	2.6
1	A	917	GLN	2.6
1	A	771	ASP	2.5
1	A	832	HIS	2.5
1	A	1007	LYS	2.5
1	A	925	VAL	2.5
1	A	1016	ASP	2.5
1	A	929	SER	2.5
1	A	1070	ASP	2.5
1	A	913	LYS	2.4
1	A	815	LEU	2.4
1	A	794	MET	2.4
1	A	1449	LEU	2.4
1	A	844	GLU	2.4
1	A	1448	ARG	2.4
1	A	787	GLU	2.4
1	A	756	ASP	2.4
1	A	800	PHE	2.4
1	A	1465	VAL	2.4
1	A	1107	ARG	2.3
1	A	842	PRO	2.3
1	A	1279	ARG	2.3
1	A	937	TYR	2.3
1	A	1006	GLY	2.2
1	A	797	ALA	2.2
1	A	826	MET	2.2
1	A	1434	PHE	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	791	GLY	2.1
1	A	777	LEU	2.1
1	A	1230	GLN	2.1
1	A	764	ASP	2.0
1	A	894	HIS	2.0

## 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](#)

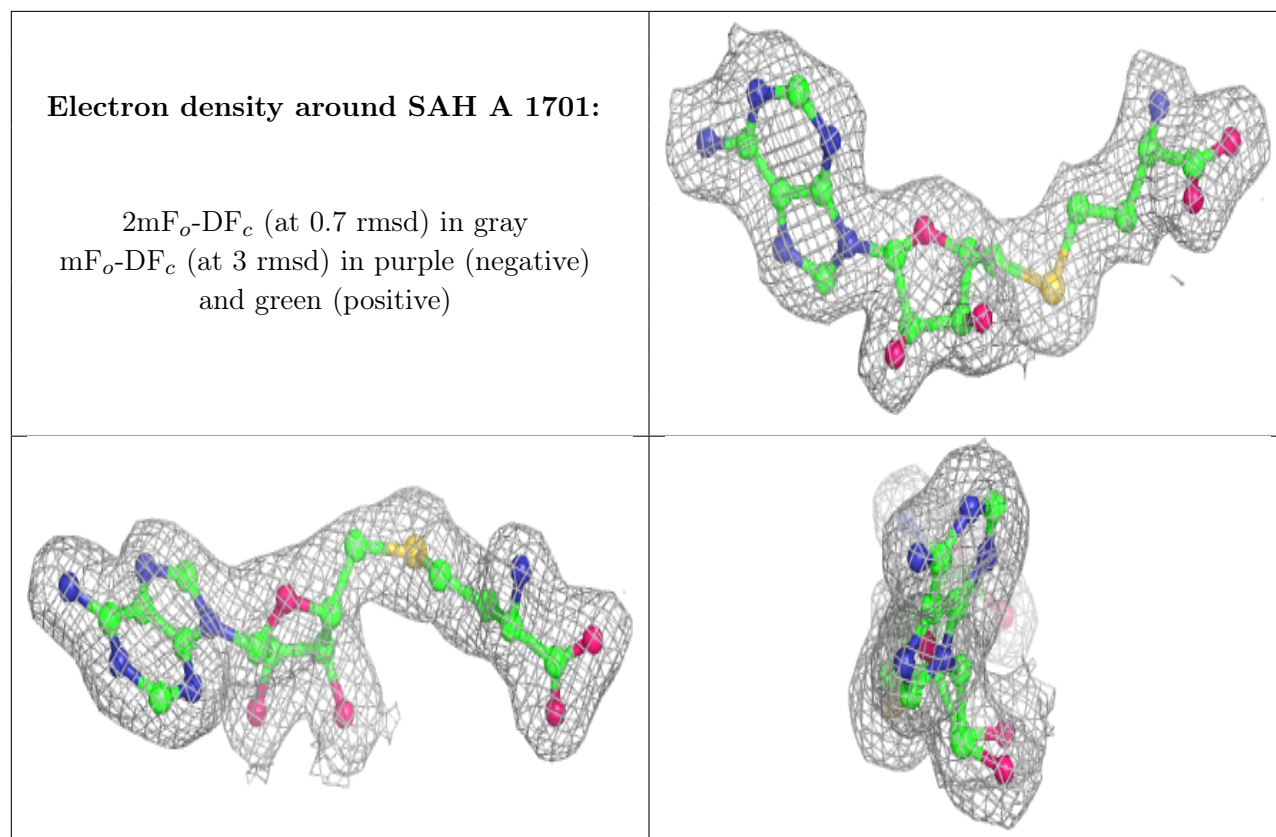
There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

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( $\text{\AA}^2$ )	Q<0.9
4	SO4	A	1712	5/5	0.81	0.40	81,83,89,95	0
4	SO4	A	1705	5/5	0.85	0.17	87,89,92,95	0
4	SO4	A	1710	5/5	0.88	0.19	75,81,85,86	0
4	SO4	A	1713	5/5	0.92	0.13	73,79,81,82	0
4	SO4	A	1707	5/5	0.93	0.21	60,60,67,79	0
4	SO4	A	1706	5/5	0.94	0.22	48,67,71,71	0
3	ZN	A	1702	1/1	0.96	0.05	50,50,50,50	0
4	SO4	A	1711	5/5	0.96	0.16	47,50,59,60	0
2	SAH	A	1701	26/26	0.98	0.11	12,15,17,20	0
4	SO4	A	1708	5/5	0.98	0.11	28,28,37,42	0
3	ZN	A	1703	1/1	0.98	0.07	25,25,25,25	0
4	SO4	A	1704	5/5	0.99	0.06	29,32,36,38	0
4	SO4	A	1709	5/5	0.99	0.09	40,44,46,49	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.