



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 07:20 PM UTC

PDB ID : 3AV5 / pdb_00003av5
Title : Crystal structure of mouse DNA methyltransferase 1 with AdoHcy
Authors : Takeshita, K.; Suetake, I.; Yamashita, E.; Suga, M.; Narita, H.; Nakagawa, A.; Tajima, S.
Deposited on : 2011-02-22
Resolution : 3.25 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

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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

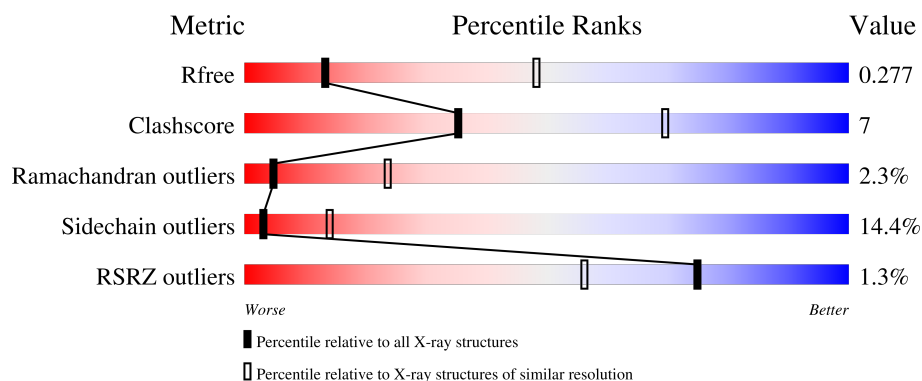
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.25 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}	180053	1605 (3.30-3.22)
Clashscore	190562	1660 (3.30-3.22)
Ramachandran outliers	187476	1630 (3.30-3.22)
Sidechain outliers	187428	1629 (3.30-3.22)
RSRZ outliers	180081	1605 (3.30-3.22)

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 ≥ 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	1330	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9135 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.

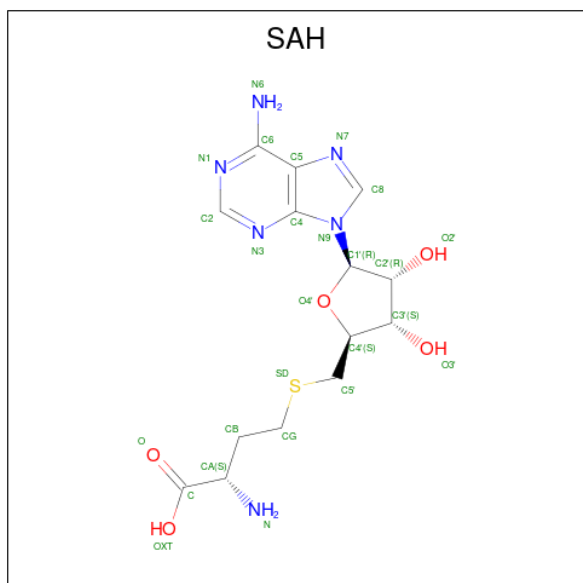
- Molecule 1 is a protein called DNA (cytosine-5)-methyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1139	Total	C	N	O	S	0	0	0
			9105	5760	1584	1701	60			

- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	4	Total	Zn	0	0
			4	4		

- Molecule 3 is S-ADENOSYL-L-HOMOCYSTEINE (CCD ID: SAH) (formula: C₁₄H₂₀N₆O₅S).

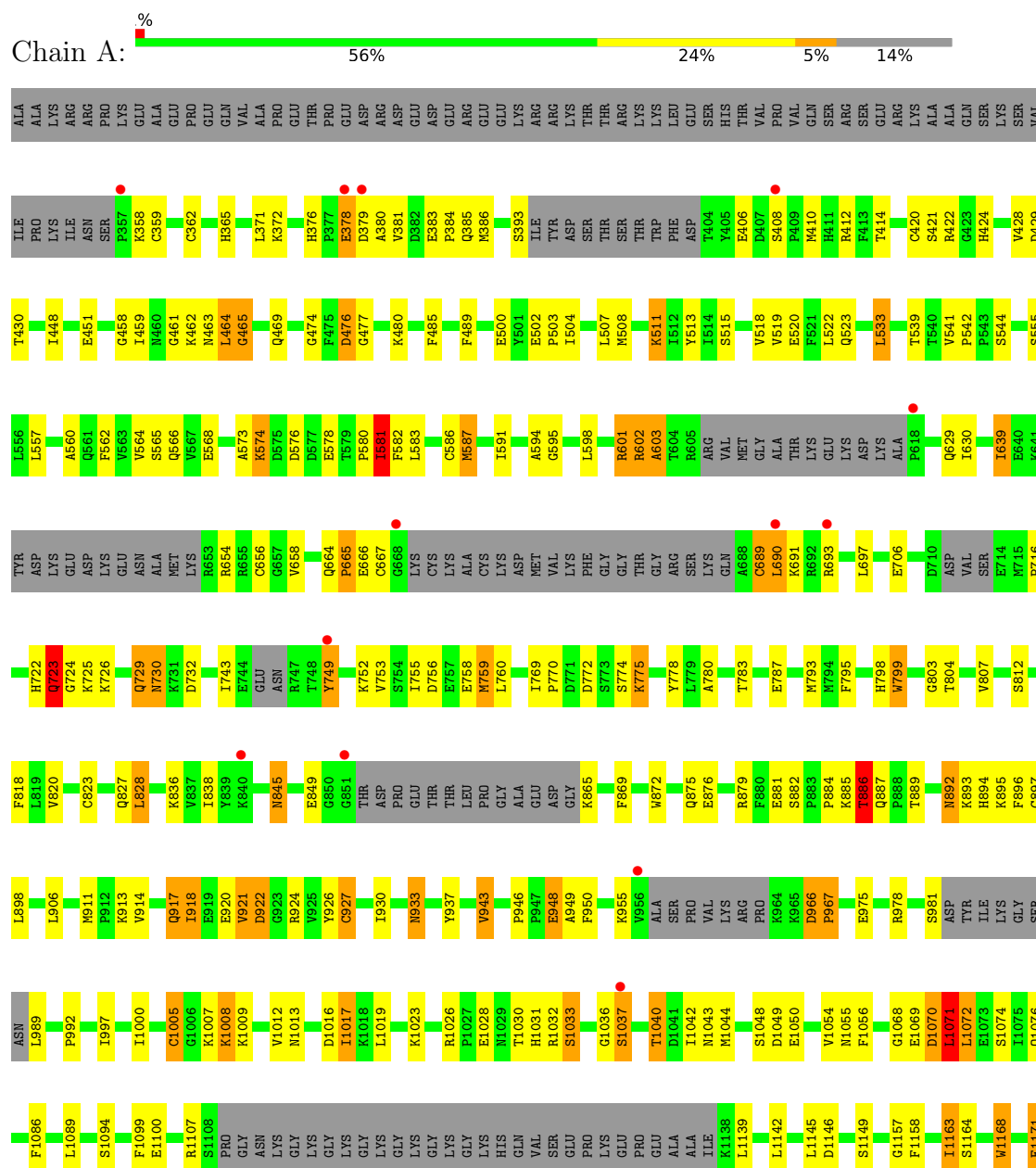


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			26	14	6	5	1		

3 Residue-property plots

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



LYS	G1517	P1410	R1292	W1172
ALA	R1518	T1411	M1293	W1173
LYS	L1519	L1412	G1294	D1174
GLU	E1520	R1413	Y1295	P1175
GLU	F1525	D1414	Q1296	A1176
ALA		H1415	L1302	A1177
ALA		I1416	V1309	Q1178
THR	T1528	C1417	A1310	A1179
LYS	V1529	D1418	Q1311	F1180
ASP	N1530	D1419	T1312	R1181
	N1531	M1420	T1311	L1182
	P1532		A1315	
	M1535	V1424		T1191
		A1425		E1192
	Q1538	A1426	I1317	
	G1539	R1427	A1320	L1197
	R1540	M1428	P1328	L1198
	V1541	R1429	L1329	
	L1542	H1430	F1330	S1209
	H1543	L1433	P1331	P1215
	P1544	S1437	E1332	Q1216
	E1545	D1438	P1333	
	Q1546	W1439		D1219
	H1547	R1440	V1336	M1222
	R1548		C1342	L1223
	V1549	D1451	Q1343	
		G1452	V1348	P1228
	V1552	V1453	D1349	G1229
	A1556	H1456	V1354	Q1230
		K1457	M1355	G1231
	I1571	T1461		F1232
	R1574		R1359	S1233
	H1575	V1465	L1360	
			S1361	R1237
	V1578		S1362	F1238
		S1470		N1239
	V1582	V1478	R1366	K1245
	P1585	P1489		L1250
	L1586	E1490	R1371	Y1257
	A1587	S1491	M1374	
	K1588	R1492	S1375	Y1261
			D1376	R1262
	L1592	I1498	L1377	P1263
	E1593	P1499	S1385	
	I1594	W1500	N1386	F1266
	C1597	C1501		L1267
	L1598	L1502		
		P1503		N1270
	S1601			V1271
		W1512		R1272
	E1604	A1513		N1273
		G1514		F1274
	A1608	L1515		V1275
	ALA	VAL		S1276
				Y1277

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	134.99Å 96.92Å 130.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.66 – 3.25 38.66 – 3.25	Depositor EDS
% Data completeness (in resolution range)	(Not available) (38.66-3.25) 95.7 (38.66-3.25)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.11 (at 3.25Å)	Xtriage
Refinement program	BUSTER-TNT, BUSTER 2.8.0	Depositor
R, R_{free}	0.183 , 0.264 0.192 , 0.277	Depositor DCC
R_{free} test set	1367 reflections (5.17%)	wwPDB-VP
Wilson B-factor (Å ²)	59.9	Xtriage
Anisotropy	0.261	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 80.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.018 for l,-k,h	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	9135	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: 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	0.89	2/9325 (0.0%)	1.43	93/12607 (0.7%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	690	LEU	CA-C	5.65	1.59	1.52
1	A	1430	HIS	C-N	5.20	1.37	1.33

The worst 5 of 93 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1360	LEU	N-CA-C	9.82	123.92	111.24
1	A	1360	LEU	CA-C-N	9.49	138.78	121.70
1	A	1360	LEU	C-N-CA	9.49	138.78	121.70
1	A	376	HIS	N-CA-C	7.56	120.72	110.29
1	A	1491	SER	N-CA-C	-7.41	104.24	113.28

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	9105	0	8877	133	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	4	0	0	0	0
3	A	26	0	19	0	0
All	All	9135	0	8896	133	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 7.

The worst 5 of 133 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:1543:HIS:HD2	1:A:1546:GLN:H	1.18	0.85
1:A:1330:PHE:H	1:A:1356:ASN:HD21	1.24	0.83
1:A:885:LYS:HA	1:A:886:THR:HG23	1.61	0.80
1:A:1068:GLY:HA2	1:A:1071:LEU:HD12	1.62	0.79
1:A:1223:LEU:HB2	1:A:1263:PRO:HG3	1.66	0.77

There are no symmetry-related clashes.

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	1117/1330 (84%)	992 (89%)	99 (9%)	26 (2%)	5 24

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	379	ASP
1	A	408	SER
1	A	465	GLY
1	A	573	ALA
1	A	896	PHE

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	1003/1162 (86%)	859 (86%)	144 (14%)	3 14

5 of 144 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1275	VAL
1	A	1604	GLU
1	A	1313	ARG
1	A	1433	LEU
1	A	759	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1356	ASN
1	A	1409	GLN
1	A	1543	HIS
1	A	739	GLN
1	A	729	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 oligosaccharides in this entry.

5.6 Ligand geometry

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 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$
3	SAH	A	1	-	27,28,28	1.17	2 (7%)	36,40,40	2.56	15 (41%)

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	-	-	3/15/31/31	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1	SAH	C6-N6	2.75	1.41	1.34
3	A	1	SAH	C2'-C3'	2.15	1.59	1.53

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1	SAH	C5-C4-N3	-5.84	118.68	126.72
3	A	1	SAH	N3-C4-N9	5.63	136.74	127.17
3	A	1	SAH	N3-C2-N1	-4.74	121.41	128.58
3	A	1	SAH	O4'-C1'-C2'	-4.56	96.85	106.62
3	A	1	SAH	C5-N7-C8	4.26	110.14	103.45

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1	SAH	CA-CB-CG-SD

Continued on next page...

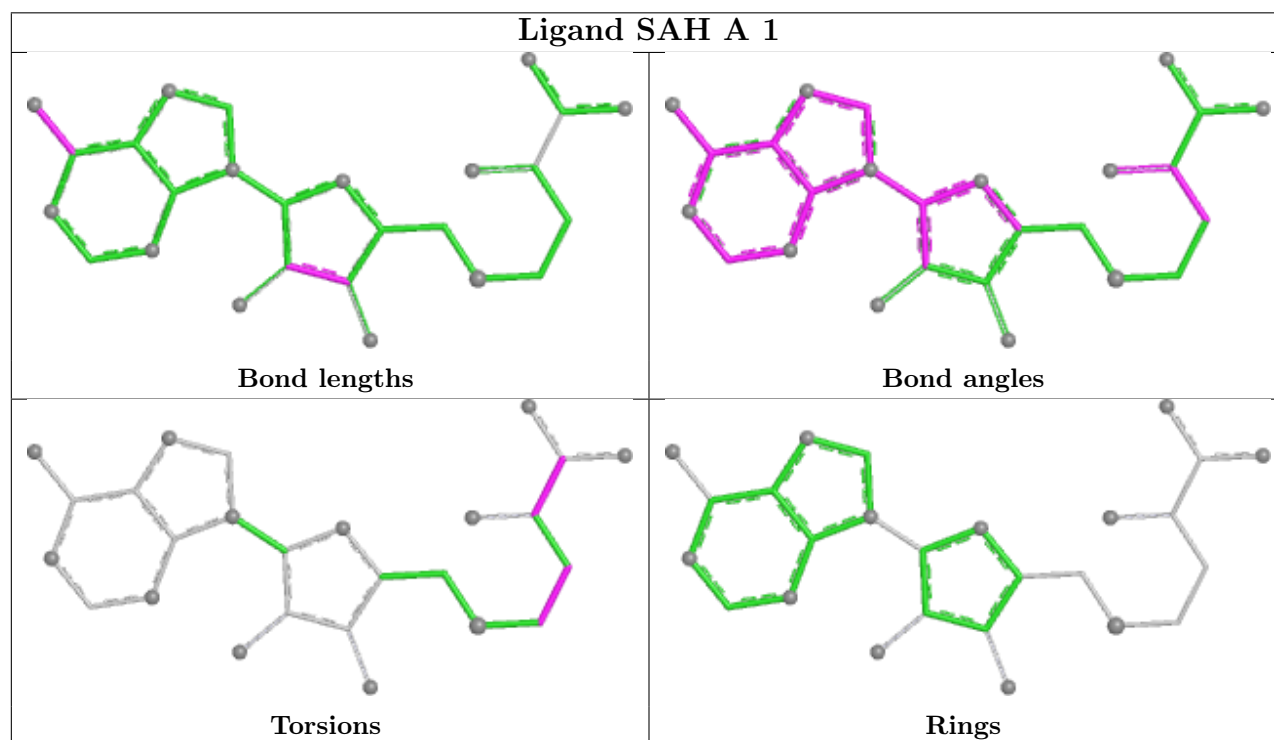
Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	A	1	SAH	O-C-CA-N
3	A	1	SAH	OXT-C-CA-N

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 ⓘ

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, 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	1139/1330 (85%)	-0.09	15 (1%) 75 56	16, 56, 103, 125	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1608	ALA	5.4
1	A	851	GLY	4.4
1	A	690	LEU	3.9
1	A	408	SER	3.1
1	A	668	GLY	3.1

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 oligosaccharides 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, 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.

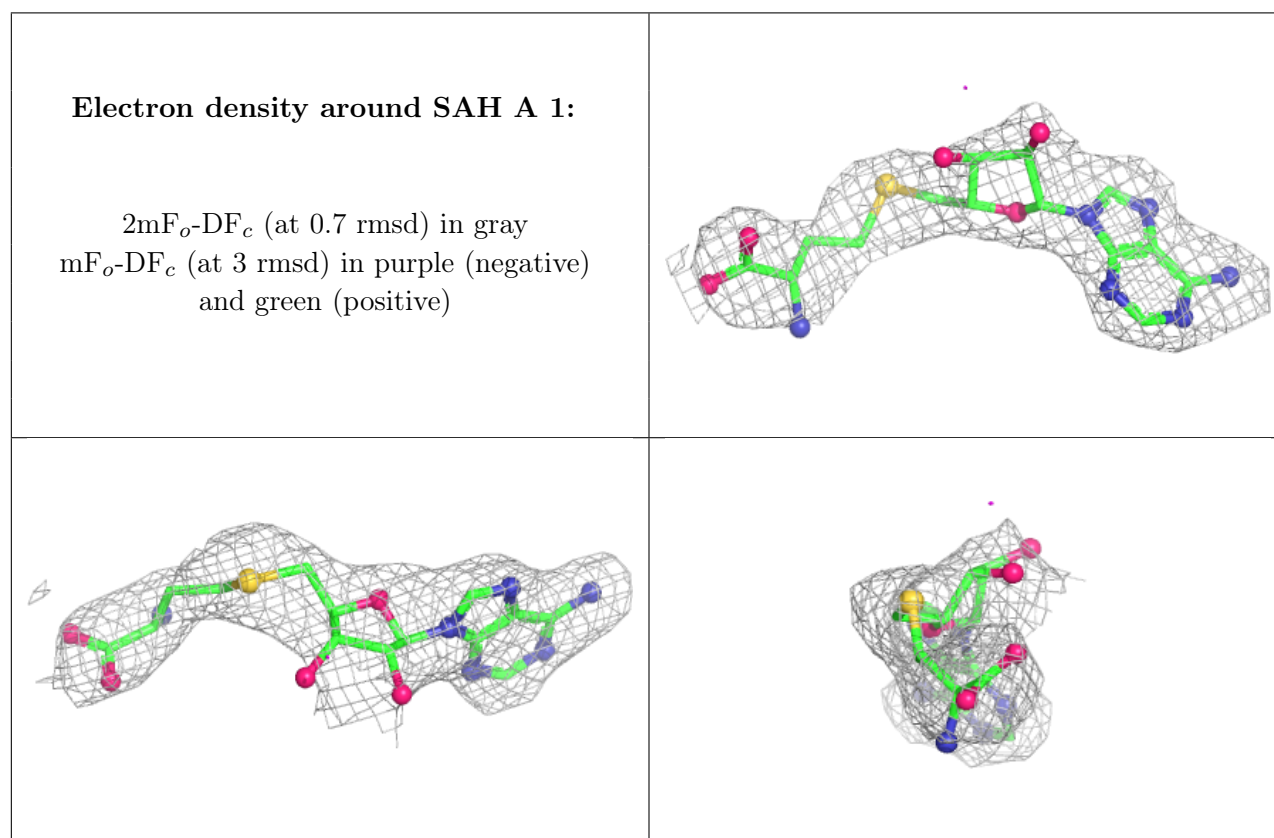
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	SAH	A	1	26/26	0.98	0.07	38,41,51,52	0
2	ZN	A	2002	1/1	1.00	0.02	65,65,65,65	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	ZN	A	2004	1/1	1.00	0.02	58,58,58,58	0
2	ZN	A	2005	1/1	1.00	0.01	34,34,34,34	0
2	ZN	A	2001	1/1	1.00	0.03	56,56,56,56	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.