



wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2020 – 04:29 am BST

PDB ID : 5DPL
Title : The structure of PKMT2 from Rickettsia typhi in complex with AdoHcy
Authors : Noinaj, N.; Abeykoon, A.; He, Y.; Yang, D.C.; Buchanan, S.K.
Deposited on : 2015-09-12
Resolution : 3.20 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 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 : 2.11
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) : 2.11

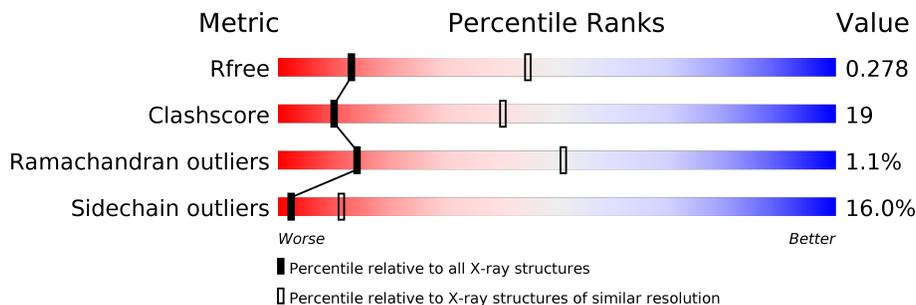
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)

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 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	535	
1	B	535	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 8077 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.

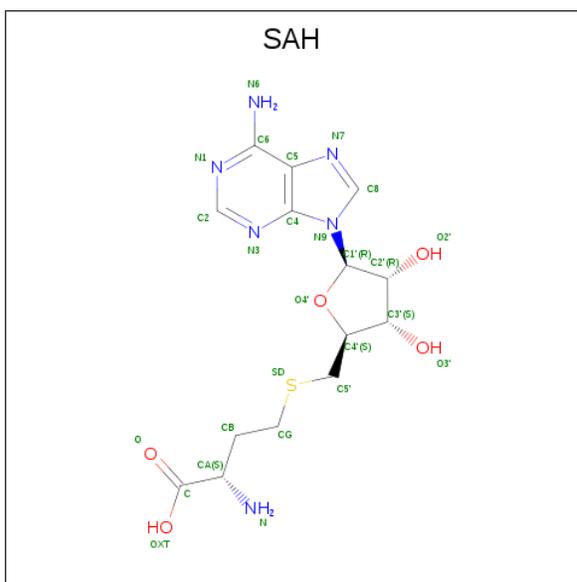
- Molecule 1 is a protein called protein lysine methyltransferase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	520	Total 4055	C 2615	N 659	O 768	S 13	0	1	0
1	B	512	Total 3970	C 2566	N 643	O 749	S 12	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP Q68XQ5
B	0	GLY	-	expression tag	UNP Q68XQ5

- Molecule 2 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: C₁₄H₂₀N₆O₅S).



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

Continued on next page...

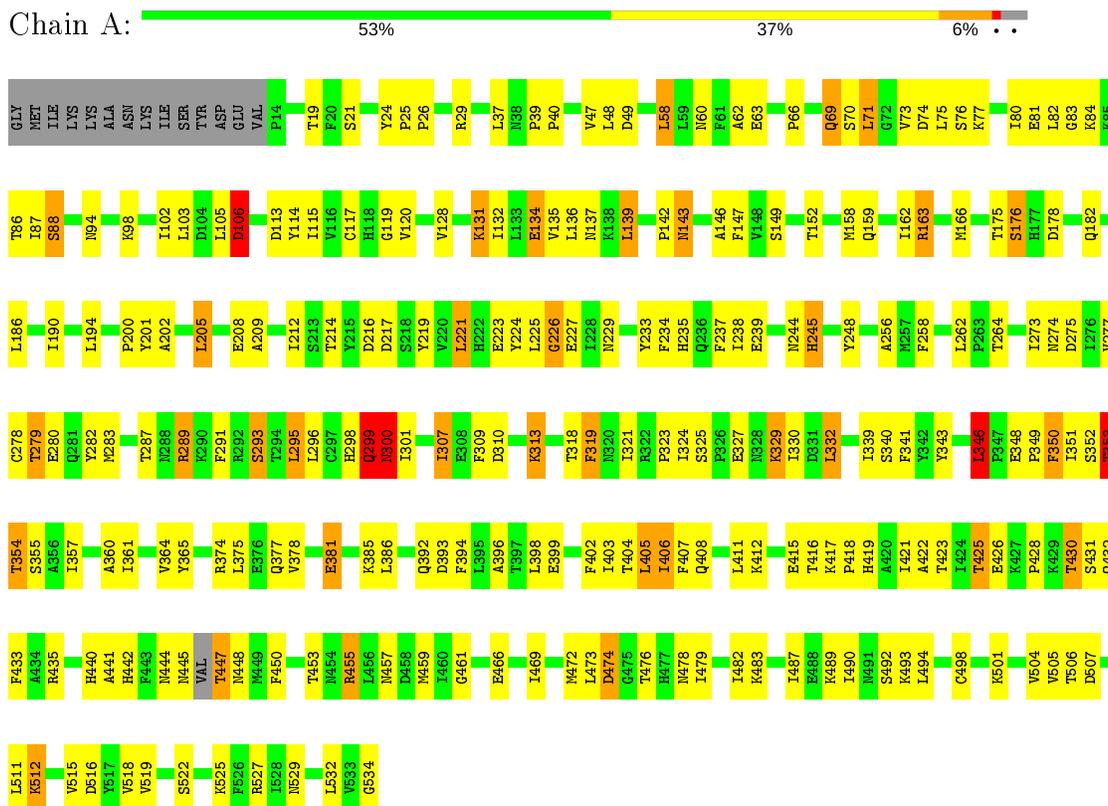
Continued from previous page...

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

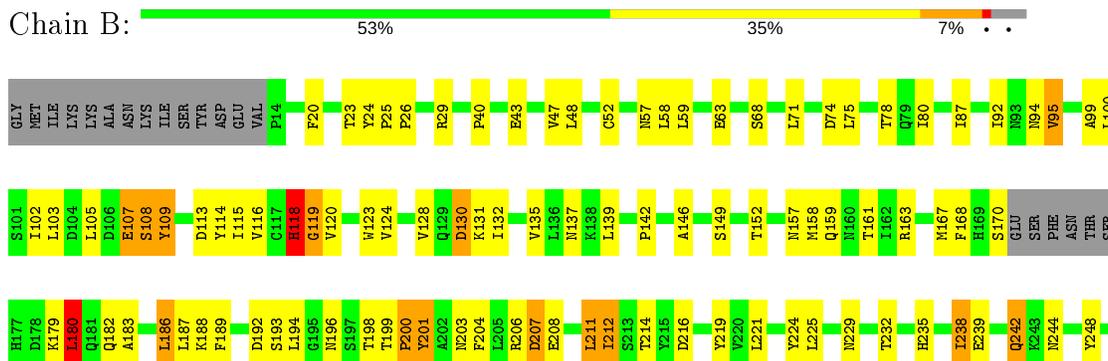
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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. 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. 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: protein lysine methyltransferase 2



- Molecule 1: protein lysine methyltransferase 2



S253	L246	T430	L511
A256	P347	S431	K512
M257	E348	Q432	E513
F258	P349	F433	F514
I259	F350	A434	V515
G260	I351	R435	V518
L270	S252	Y436	
I273	I353	Q437	
N274	S355	K439	V521
C278	A356	F443	S522
M283	I357	ASN	L523
D284	I361	ASN	E524
F285	L362	VAL	R527
L286	Y363	T447	L531
T287	V364	F450	L532
M288	Y365		V533
R289	M368	T453	G534
K290	I369	M454	
F291	S370	R455	
R292	N371	I462	
L295	P372	H465	
L296	L375	E466	
C297	L386	I469	
H298	Y389	M472	
N304	R390	L473	
R305	L391	D474	
K306	Q392	H477	
I307	D393	N478	
L312	L398	I479	
K313	E399	D480	
Y316	Q400	D481	
N320	H401	I482	
I321	F402	K483	
S325	T403	I487	
I330	L404	E488	
D331	L405	N491	
L332	I406	S492	
N333	F407	K493	
E337	Q408	L494	
N338	L411	L495	
I339	T416	D499	
S340	H419	V504	
F341	A422	V505	
Y342	T425	T506	
I343	E426	D507	
E344	K427	P508	
N345	P428	K509	
	K429	L510	

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	83.35Å 91.03Å 105.83Å 90.00° 112.30° 90.00°	Depositor
Resolution (Å)	48.95 – 3.20 48.96 – 3.19	Depositor EDS
% Data completeness (in resolution range)	97.4 (48.95-3.20) 93.5 (48.96-3.19)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.39 (at 3.19Å)	Xtrriage
Refinement program	PHENIX (1.10_2142)	Depositor
R, R_{free}	0.244 , 0.280 0.245 , 0.278	Depositor DCC
R_{free} test set	2000 reflections (8.36%)	wwPDB-VP
Wilson B-factor (Å ²)	68.7	Xtrriage
Anisotropy	0.153	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , -13.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.24$	Xtrriage
Estimated twinning fraction	0.080 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	8077	wwPDB-VP
Average B, all atoms (Å ²)	72.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.54% 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

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	0.32	0/4145	0.62	3/5643 (0.1%)
1	B	0.32	0/4059	0.63	2/5530 (0.0%)
All	All	0.32	0/8204	0.63	5/11173 (0.0%)

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	6
1	B	0	5
All	All	0	11

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	346	LEU	CA-CB-CG	7.28	132.05	115.30
1	B	118	HIS	C-N-CA	-6.63	108.38	122.30
1	B	180	LEU	CA-CB-CG	6.23	129.63	115.30
1	A	205	LEU	CA-CB-CG	5.06	126.95	115.30
1	A	353	THR	C-N-CA	5.02	134.25	121.70

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	106	ASP	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	176	SER	Peptide
1	A	226	GLY	Peptide
1	A	299	GLN	Peptide
1	A	501	LYS	Peptide

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	4055	0	3881	151	0
1	B	3970	0	3781	151	0
2	A	26	0	19	2	0
2	B	26	0	19	2	0
All	All	8077	0	7700	300	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 19.

The worst 5 of 300 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:B:320:ASN:HA	1:B:369:ILE:HD11	1.36	1.03
1:A:435:ARG:NH1	1:A:473:LEU:O	2.01	0.92
1:B:186:LEU:O	1:B:189:PHE:HB3	1.68	0.92
1:B:393:ASP:OD1	1:B:393:ASP:N	2.05	0.89
1:A:421:ILE:HG22	1:A:423:THR:H	1.42	0.84

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	517/535 (97%)	445 (86%)	67 (13%)	5 (1%)	15	54
1	B	507/535 (95%)	443 (87%)	58 (11%)	6 (1%)	13	49
All	All	1024/1070 (96%)	888 (87%)	125 (12%)	11 (1%)	14	51

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	300	ASN
1	A	504	VAL
1	B	505	VAL
1	A	299	GLN
1	B	200	PRO

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	426/484 (88%)	358 (84%)	68 (16%)	2	11
1	B	414/484 (86%)	348 (84%)	66 (16%)	2	12
All	All	840/968 (87%)	706 (84%)	134 (16%)	2	11

5 of 134 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	440	HIS
1	B	100	LEU
1	B	425	THR
1	A	453	THR
1	B	24	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	69	GLN
1	A	137	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

2 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
2	SAH	A	601	-	21,28,28	1.19	2 (9%)	20,40,40	1.72	3 (15%)
2	SAH	B	601	-	21,28,28	1.19	2 (9%)	20,40,40	1.78	3 (15%)

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	601	-	-	3/7/31/31	0/3/3/3
2	SAH	B	601	-	-	3/7/31/31	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	601	SAH	C2-N3	4.02	1.38	1.32
2	A	601	SAH	C2-N3	4.01	1.38	1.32
2	A	601	SAH	C2-N1	2.45	1.38	1.33
2	B	601	SAH	C2-N1	2.41	1.38	1.33

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	601	SAH	N3-C2-N1	-5.37	120.29	128.68
2	A	601	SAH	N3-C2-N1	-5.30	120.39	128.68
2	B	601	SAH	C5'-SD-CG	-4.33	89.29	102.27
2	A	601	SAH	C5'-SD-CG	-3.68	91.22	102.27
2	B	601	SAH	C3'-C2'-C1'	2.47	104.70	100.98

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	601	SAH	N-CA-CB-CG
2	A	601	SAH	C-CA-CB-CG
2	B	601	SAH	N-CA-CB-CG
2	B	601	SAH	C-CA-CB-CG
2	B	601	SAH	CB-CG-SD-C5'

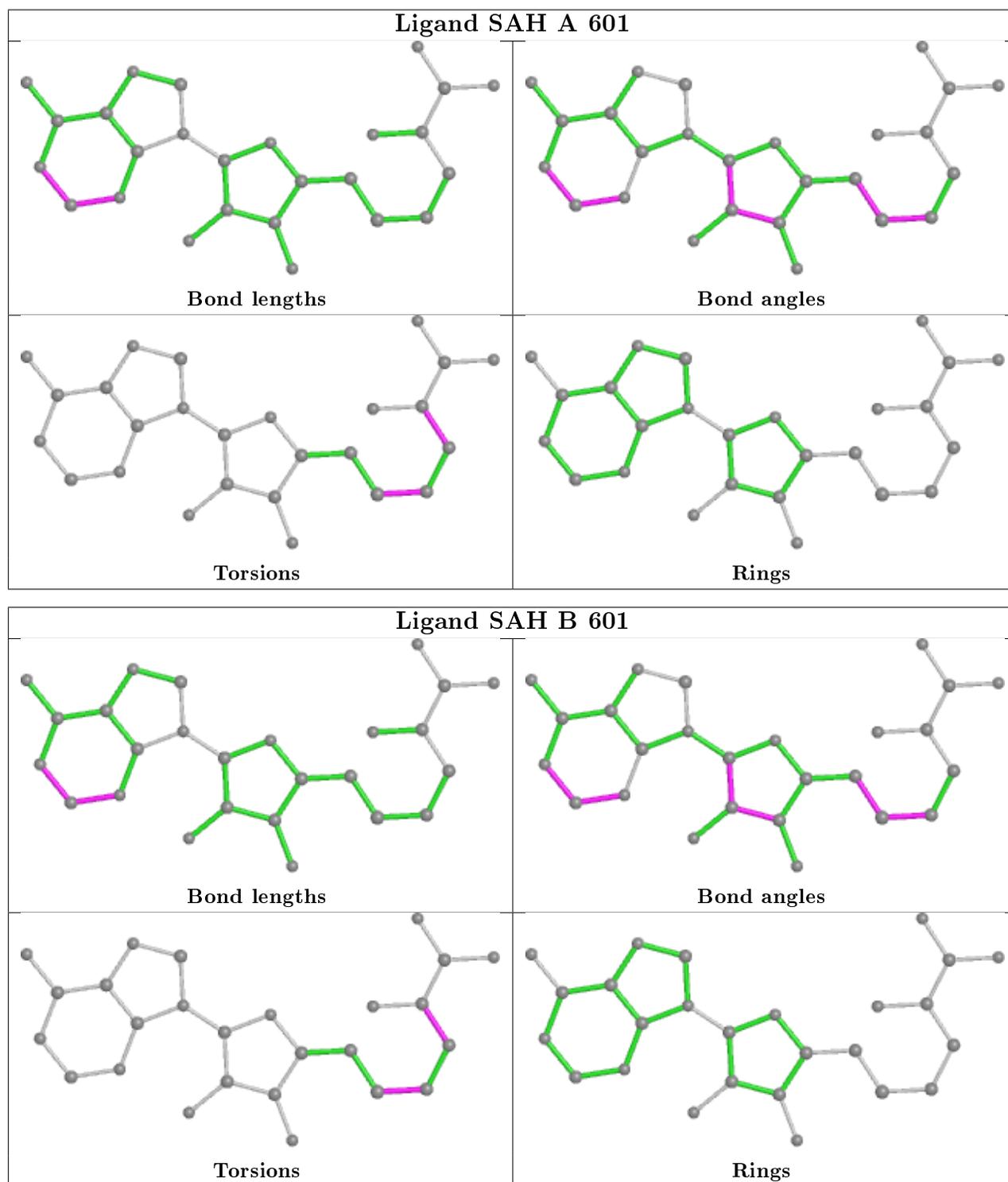
There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	SAH	2	0
2	B	601	SAH	2	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

There are no chain breaks in this entry.

6 Fit of model and data [\(i\)](#)

6.1 Protein, DNA and RNA chains [\(i\)](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

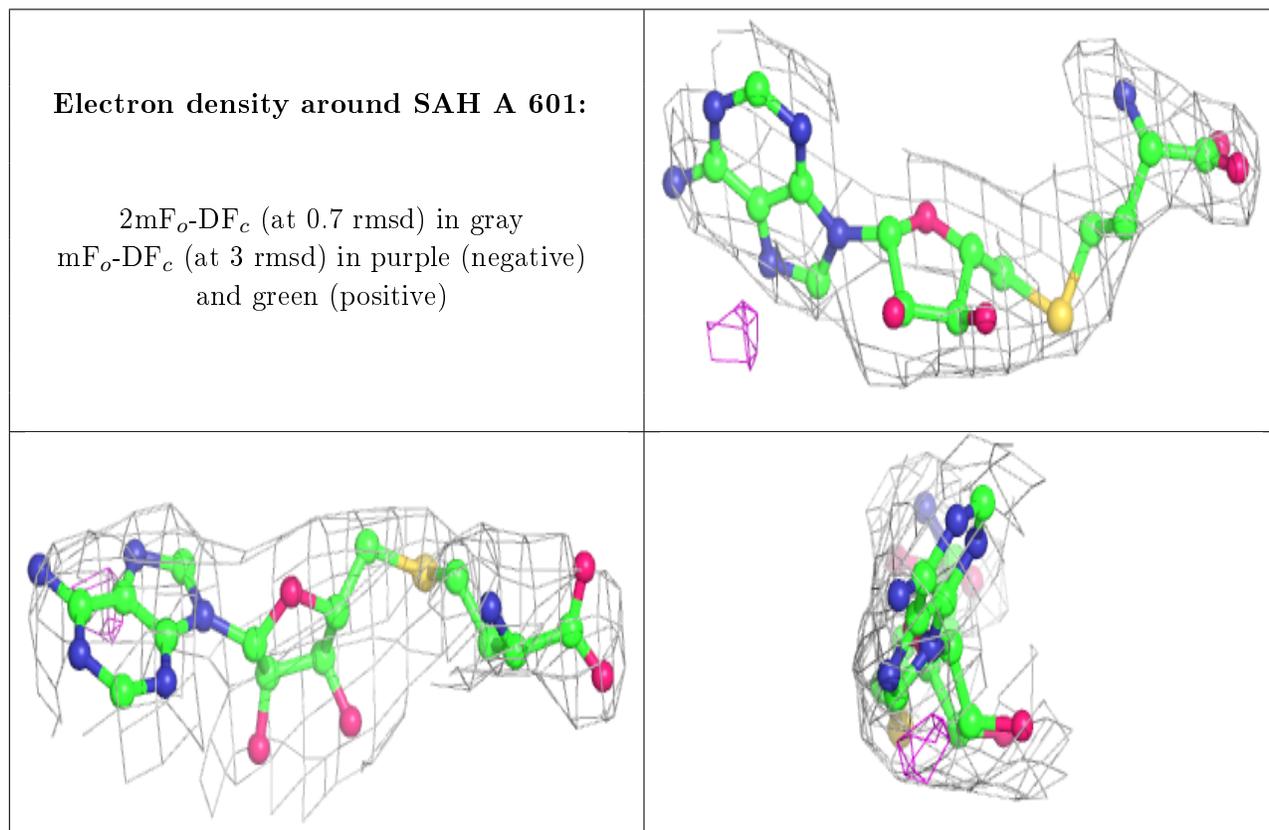
6.3 Carbohydrates [\(i\)](#)

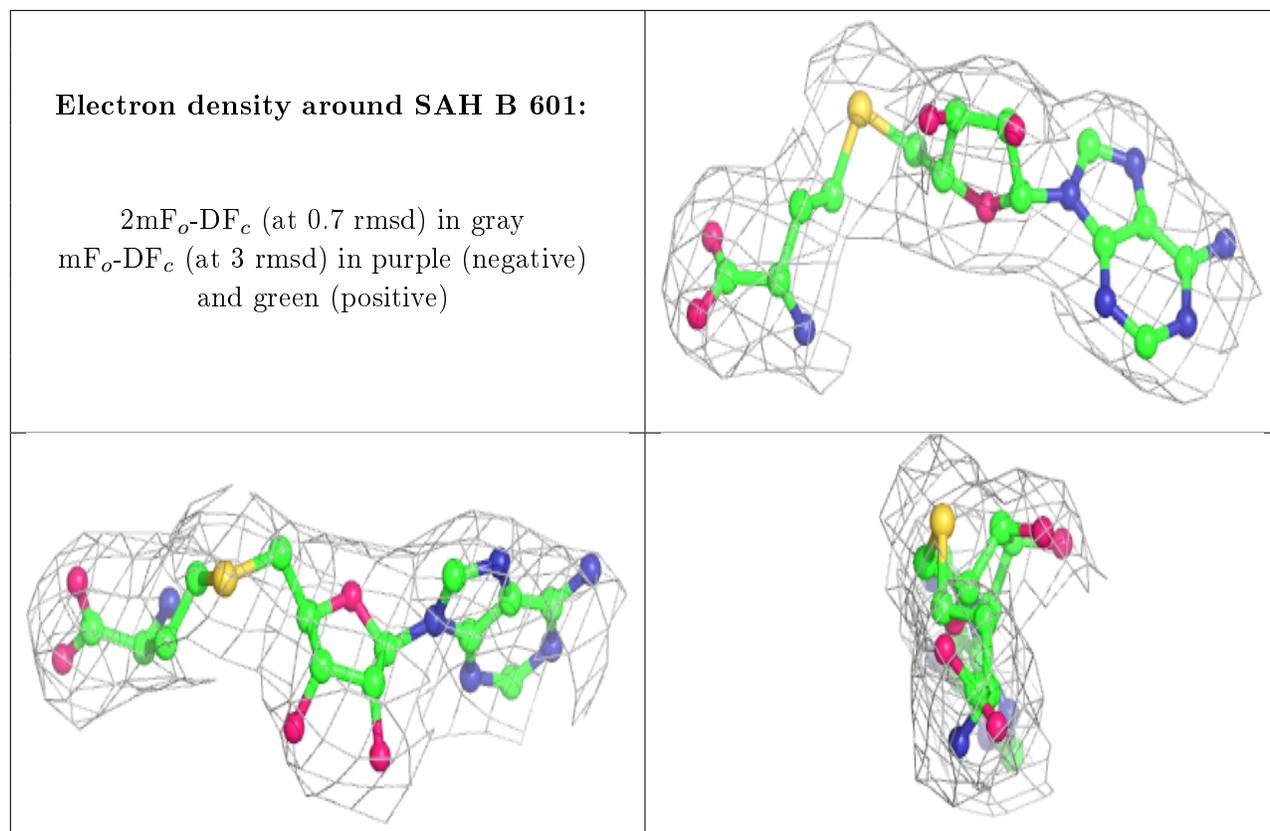
Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands [\(i\)](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

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](#)

Unable to reproduce the depositors R factor - this section is therefore empty.