



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 06:31 am BST

PDB ID : 5DPD
Title : The structure of PKMT1 from Rickettsia prowazekii in complex with AdoMet
Authors : Noinaj, N.; Abeykoon, A.; He, Y.; Yang, D.C.; Buchanan, S.K.
Deposited on : 2015-09-12
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

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

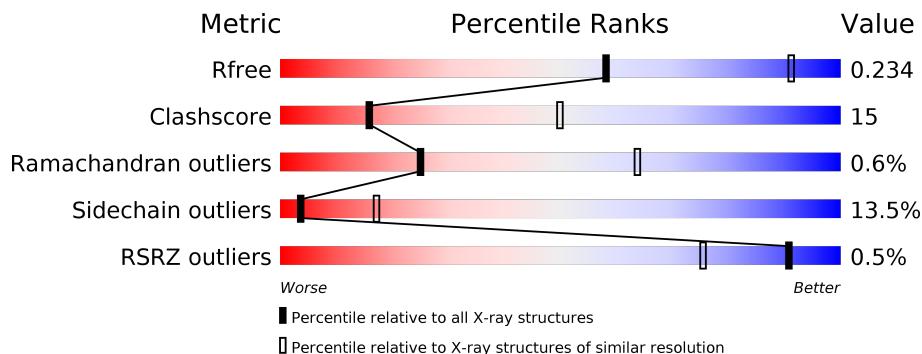
1 Overall quality at a glance

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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (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 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 $\leq 5\%$. 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	554	 57% 30% 5% 8%
1	B	554	 58% 29% 6% 8%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 8141 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.

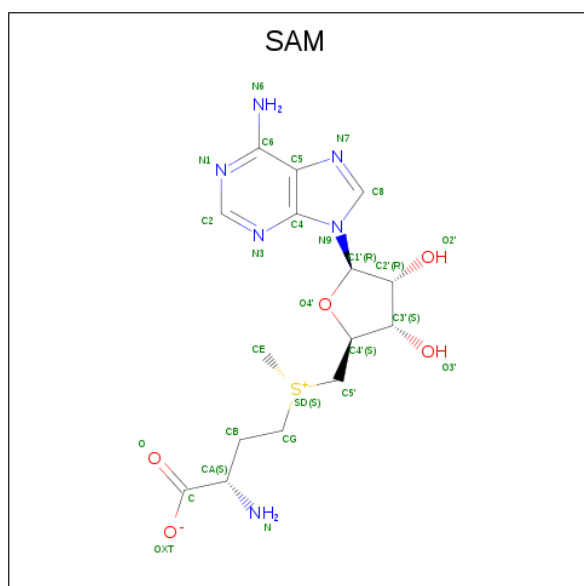
- Molecule 1 is a protein called protein lysine methyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	511	4071	2592	699	760	20	0	1	0
1	B	511	4016	2557	689	750	20	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 O05979
B	0	GLY	-	expression tag	UNP O05979

- Molecule 2 is S-ADENOSYLMETHIONINE (three-letter code: SAM) (formula: C₁₅H₂₂N₆O₅S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	A	1	27	15	6	5	1	0	0

Continued on next page...

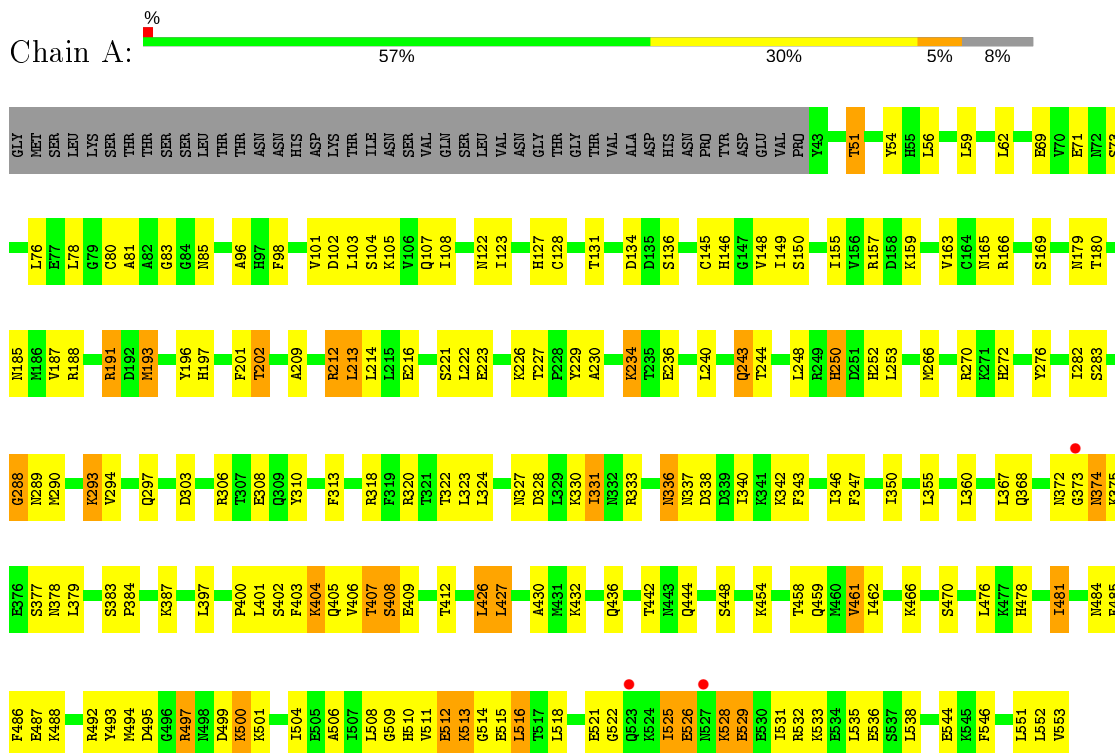
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	B	1	27	15	6	5	1	0	0

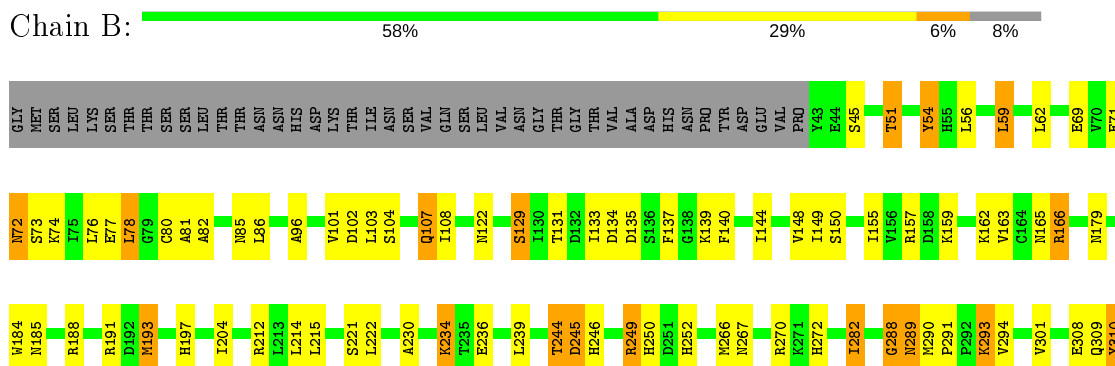
3 Residue-property plots [i](#)

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 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: protein lysine methyltransferase 1



- Molecule 1: protein lysine methyltransferase 1



M811	T392	R492
D812	F393	Y493
F313	S394	M494
	E395	
N316	N396	N498
R317		D499
R318	P400	K500
F319	L401	K501
R320	S402	A502
L323	F403	I503
L324	K404	I504
	Q405	E505
	Y406	A506
D828	T407	I507
	S408	L508
I331	E409	G509
R332		H510
R333	L419	V511
N334		E512
I335	L426	K513
N336	L427	G514
		E515
D839	A430	L516
I340	M431	T517
K341		I518
K342	T442	S519
F343	M443	R520
	Q444	E521
I346	G522	G522
F347	S448	Q523
I350	T458	E526
P351		M527
	V461	K528
L355		
T364	Q464	L635
	A465	F639
K466	K466	
Y467	Y467	
T468	T468	E544
L371	P469	K545
N372	S470	F546
G373	M471	
N374		L551
K375	M475	L552
E376	L476	V553
S377	K477	
N378	H478	
L379	E479	
S380	P480	
T381	I481	
T382		
	M484	
Y385	F485	
K386	F486	
K387	F487	
A388	K488	
I389	F489	
L390	A490	
Y391	L491	

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	98.03Å 62.94Å 107.63Å 90.00° 100.42° 90.00°	Depositor
Resolution (Å)	48.21 – 3.00 48.21 – 2.99	Depositor EDS
% Data completeness (in resolution range)	94.7 (48.21-3.00) 92.4 (48.21-2.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.65 (at 3.01Å)	Xtrriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.197 , 0.234 0.199 , 0.234	Depositor DCC
R_{free} test set	2000 reflections (7.95%)	wwPDB-VP
Wilson B-factor (Å ²)	52.0	Xtrriage
Anisotropy	0.317	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.23 , 23.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	8141	wwPDB-VP
Average B, all atoms (Å ²)	60.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.64% 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: SAM

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	0.40	0/4155	0.67	1/5633 (0.0%)
1	B	0.38	0/4100	0.63	1/5568 (0.0%)
All	All	0.39	0/8255	0.65	2/11201 (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	3
1	B	0	3
All	All	0	6

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	213	LEU	CA-CB-CG	7.47	132.49	115.30
1	B	288	GLY	N-CA-C	5.60	127.11	113.10

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	288	GLY	Peptide
1	A	374	ASN	Peptide
1	A	525	ILE	Peptide
1	B	244	THR	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	310[B]	TYR	Mainchain

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	4071	0	4007	118	0
1	B	4016	0	3892	125	0
2	A	27	0	22	5	0
2	B	27	0	22	2	0
All	All	8141	0	7943	235	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 15.

The worst 5 of 235 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:310[B]:TYR:O	1:B:311:MET:C	1.86	1.07
1:B:310[B]:TYR:O	1:B:312:ASP:N	1.91	1.03
1:A:165:ASN:ND2	1:A:272:HIS:O	1.97	0.96
1:B:212:ARG:NH2	1:B:245:ASP:OD1	1.98	0.95
1:B:165:ASN:ND2	1:B:272:HIS:O	2.03	0.92

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	510/554 (92%)	473 (93%)	36 (7%)	1 (0%)	47	82
1	B	510/554 (92%)	470 (92%)	35 (7%)	5 (1%)	15	53
All	All	1020/1108 (92%)	943 (92%)	71 (7%)	6 (1%)	25	64

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	485	PHE
1	B	527	ASN
1	A	485	PHE
1	B	289	ASN
1	B	245	ASP

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	446/504 (88%)	384 (86%)	62 (14%)	3	16
1	B	431/504 (86%)	375 (87%)	56 (13%)	4	19
All	All	877/1008 (87%)	759 (86%)	118 (14%)	4	17

5 of 118 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	512	GLU
1	B	76	LEU
1	B	471	MET
1	A	513	LYS
1	A	538	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	478	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	326	HIS
1	B	210	GLN
1	A	368	GLN
1	B	197	HIS

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	SAM	B	601	-	21,29,29	1.27	3 (14%)	18,42,42	2.05	4 (22%)
2	SAM	A	601	-	21,29,29	1.33	3 (14%)	18,42,42	2.45	5 (27%)

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	SAM	B	601	-	-	1/8/33/33	0/3/3/3
2	SAM	A	601	-	-	4/8/33/33	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)
2	B	601	SAM	C2-N3	3.82	1.38	1.32
2	A	601	SAM	C2-N3	3.67	1.38	1.32
2	A	601	SAM	C2-N1	2.73	1.39	1.33
2	A	601	SAM	O4'-C4'	-2.69	1.39	1.45
2	B	601	SAM	C2-N1	2.58	1.38	1.33

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	601	SAM	C5'-SD-CG	-6.07	87.91	103.40
2	B	601	SAM	N3-C2-N1	-5.25	120.47	128.68
2	A	601	SAM	N3-C2-N1	-5.21	120.53	128.68
2	A	601	SAM	C3'-C2'-C1'	4.01	107.01	100.98
2	B	601	SAM	O4'-C4'-C5'	-3.91	99.00	108.88

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	601	SAM	N-CA-CB-CG
2	A	601	SAM	C-CA-CB-CG
2	A	601	SAM	O4'-C4'-C5'-SD
2	A	601	SAM	C3'-C4'-C5'-SD
2	B	601	SAM	CB-CG-SD-C5'

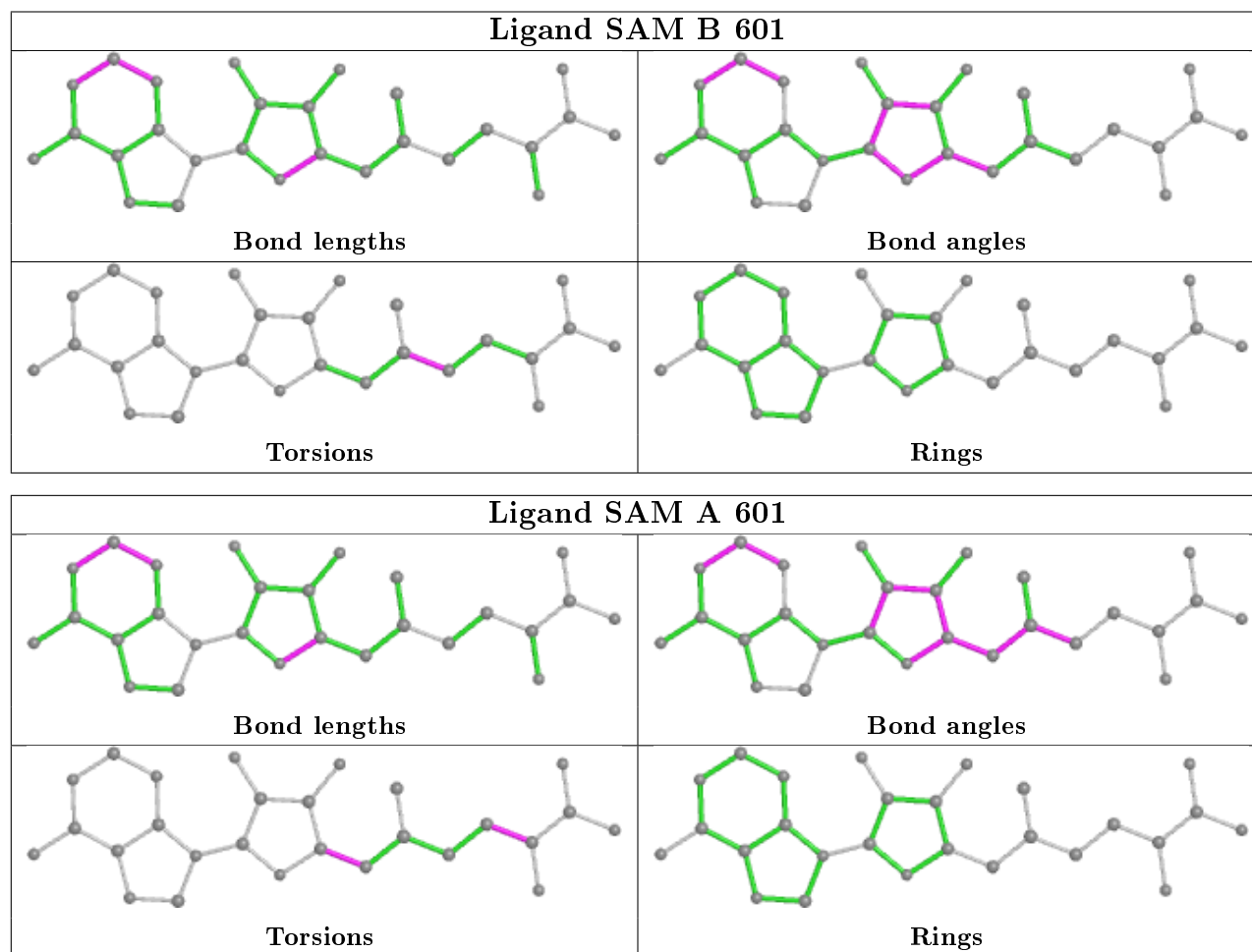
There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	601	SAM	2	0
2	A	601	SAM	5	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 [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	511/554 (92%)	-0.57	3 (0%) 89 72	22, 48, 99, 164	0
1	B	511/554 (92%)	-0.45	2 (0%) 92 79	31, 61, 122, 172	0
All	All	1022/1108 (92%)	-0.51	5 (0%) 91 75	22, 55, 111, 172	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	527	ASN	4.5
1	B	523	GLN	2.8
1	A	527	ASN	2.3
1	A	523	GLN	2.3
1	A	373	GLY	2.2

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](#)

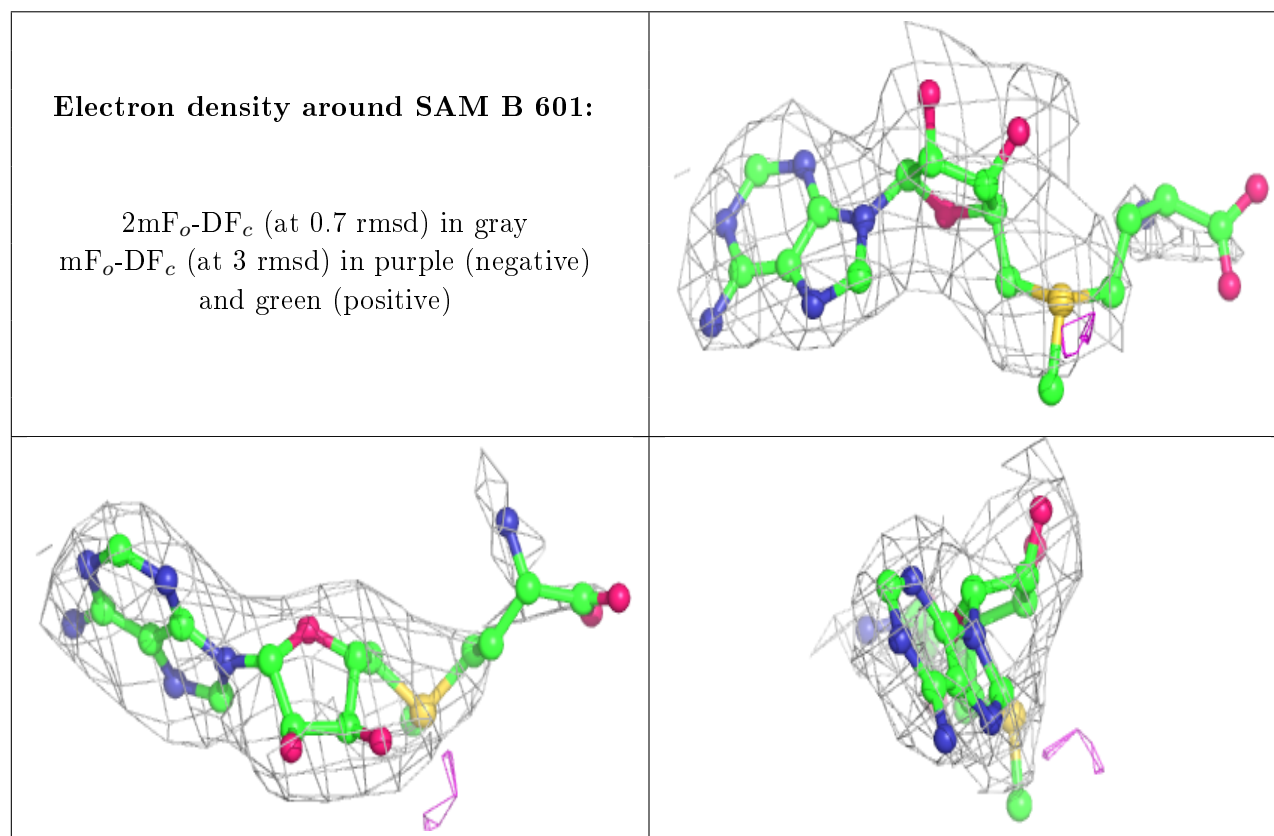
There are no carbohydrates in this entry.

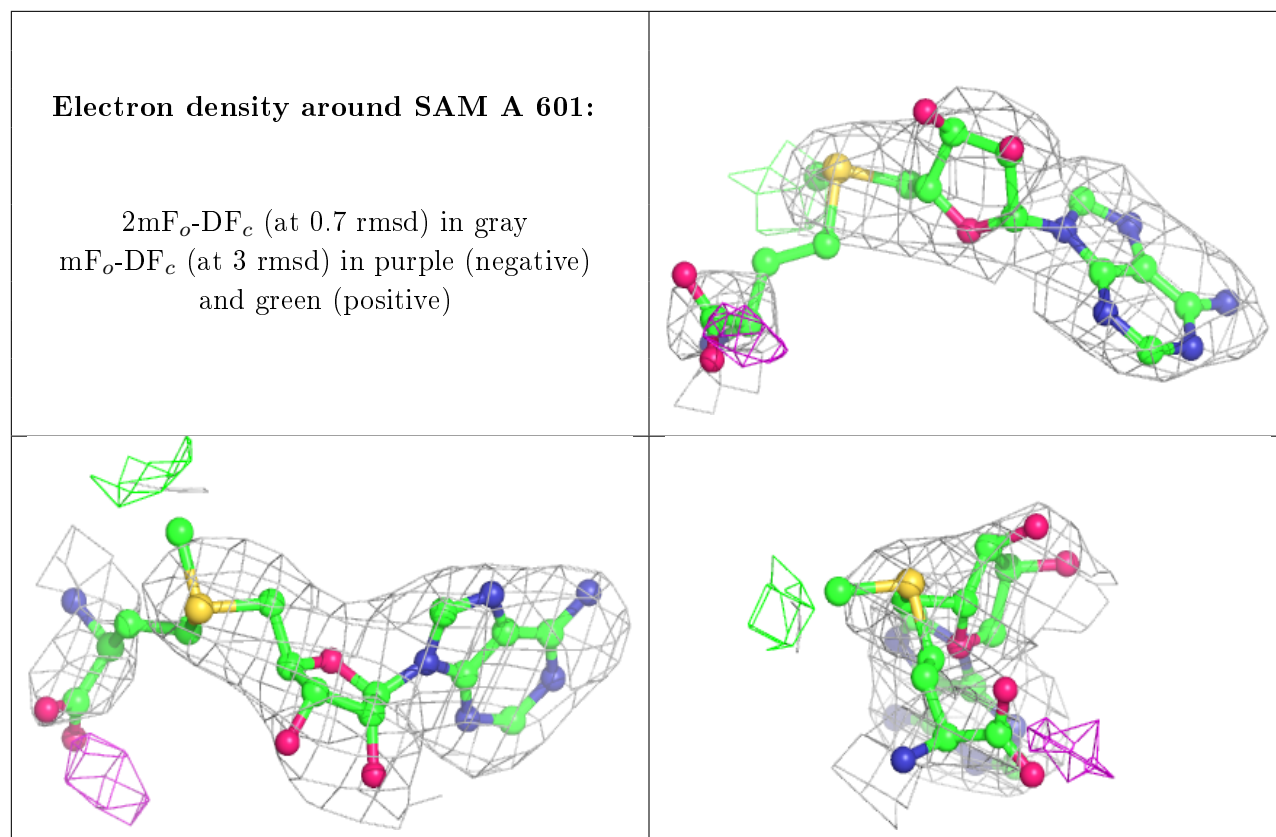
6.4 Ligands [i](#)

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(\AA^2)	Q<0.9
2	SAM	B	601	27/27	0.92	0.20	43,61,90,96	0
2	SAM	A	601	27/27	0.93	0.23	31,56,76,83	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.