Full wwPDB X-ray Structure Validation Report (i)

May 14, 2020 – 06:51 pm BST

PDB ID	:	5N0S
Title	:	Crystal structure of OphA-DeltaC6 mutant Y98A in complex with SAM
Authors	:	Song, H.; Naismith, J.H.
Deposited on	:	2017-02-03
Resolution	:	1.95 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

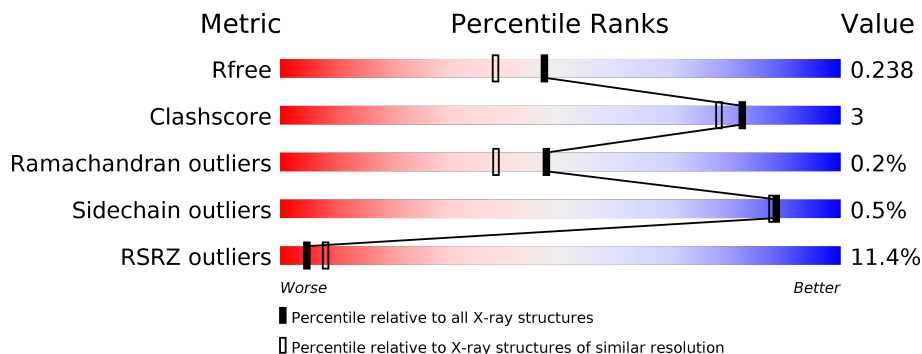|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 1.95 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	Percentile Ranks	Value
Rfree		0.238
Clashscore		3
Ramachandran outliers		0.2%
Sidechain outliers		0.5%
RSRZ outliers		11.4%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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 <=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	410	11% ... 94%
1	B	410	11% ... 94%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	MLI	B	502	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6895 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 Peptide N-methyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	401	Total	C	N	O	S	0	6	0
			3146	1996	540	591	19			
1	B	404	Total	C	N	O	S	0	5	0
			3157	2004	542	593	18			

- Molecule 2 is S-ADENOSYLMETHIONINE (three-letter code: SAM) (formula: $C_{15}H_{22}N_6O_5S$).
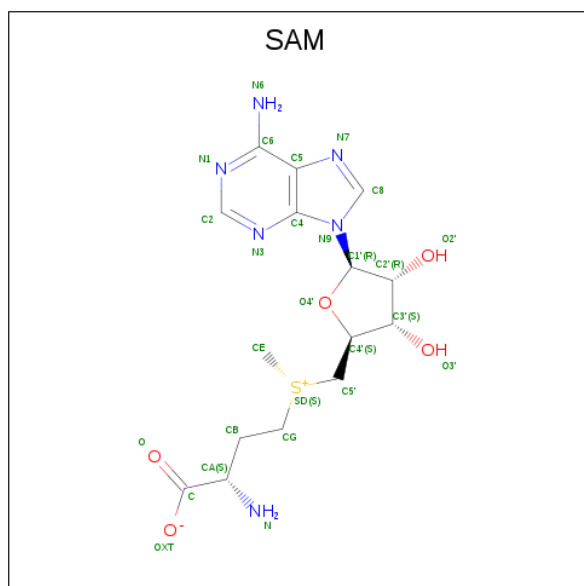
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			27	15	6	5	1		
2	B	1	Total	C	N	O	S	0	0
			27	15	6	5	1		

- Molecule 3 is MALONATE ION (three-letter code: MLI) (formula: $C_3H_2O_4$).

MLI
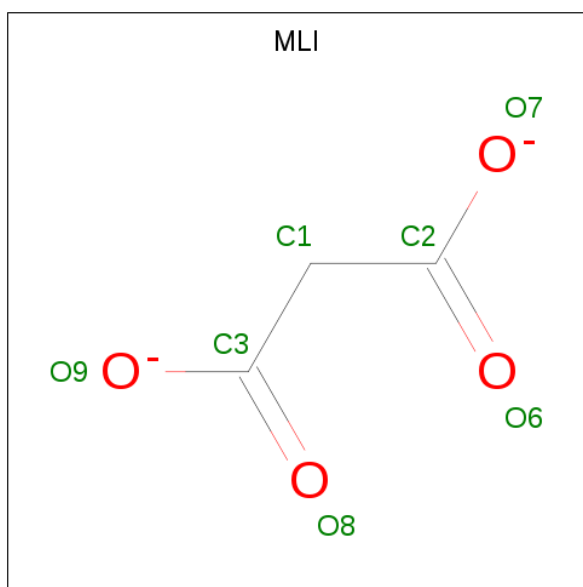
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 7	C 3	O 4	0	0
3	B	1	Total 7	C 3	O 4	0	0

• Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	259	Total 259	O 259	0	0
4	B	265	Total 265	O 265	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 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: Peptide N-methyltransferase
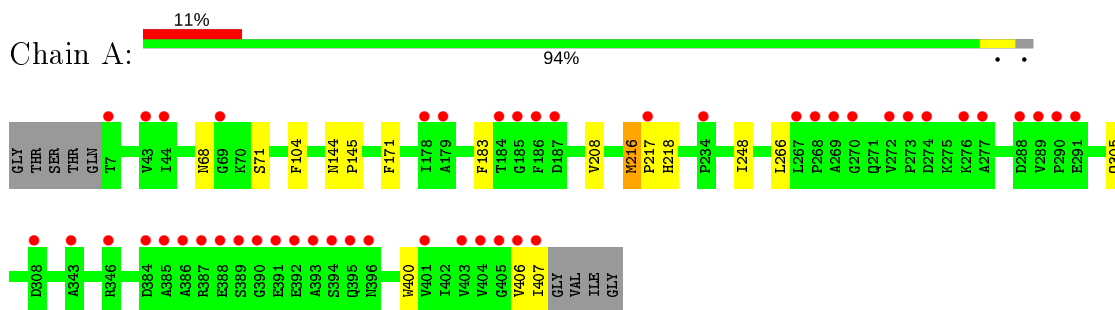
- Molecule 1: Peptide N-methyltransferase
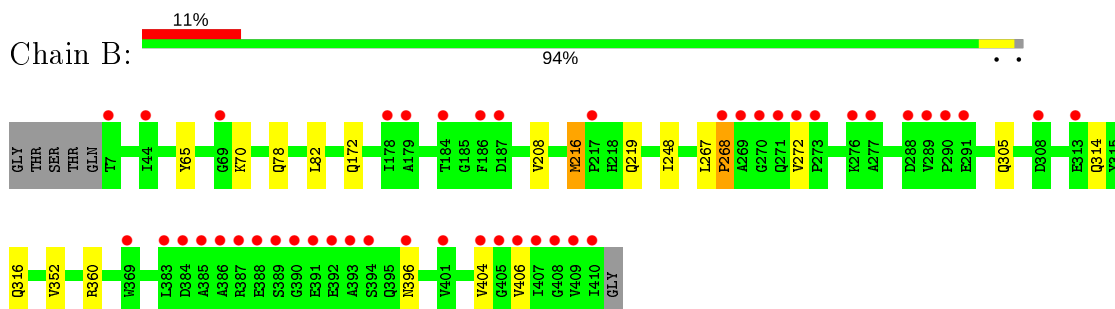
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	163.91Å 92.52Å 85.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	85.33 – 1.95 30.80 – 1.95	Depositor EDS
% Data completeness (in resolution range)	98.3 (85.33-1.95) 98.4 (30.80-1.95)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.56 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.205 , 0.233 0.211 , 0.238	Depositor DCC
R_{free} test set	4632 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	35.8	Xtriage
Anisotropy	0.117	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 50.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.51, < L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6895	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 82.87 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.5074e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1] Intensities estimated from amplitudes.
[2] Theoretical values of $< \|L\| >$, $< L^2 >$ 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: MLI, 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	0.45	0/3223	0.67	0/4384
1	B	0.46	0/3234	0.68	0/4400
All	All	0.45	0/6457	0.68	0/8784

There are no bond length outliers.

There are no bond angle outliers.

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	3146	0	3112	19	0
1	B	3157	0	3127	16	0
2	A	27	0	22	0	0
2	B	27	0	22	0	0
3	A	7	0	2	1	0
3	B	7	0	2	4	0
4	A	259	0	0	0	0
4	B	265	0	0	0	0
All	All	6895	0	6287	32	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 3.

All (32) 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:216[B]:MET:HB3	1:A:217[B]:PRO:HD2	1.35	1.08
1:A:216[B]:MET:HB3	1:A:217[B]:PRO:CD	1.84	1.03
1:B:216[A]:MET:HG3	1:B:219[A]:GLN:CD	2.15	0.67
1:A:208:VAL:HG22	1:A:248:ILE:HD12	1.78	0.64
1:B:404:VAL:HG12	1:B:406:VAL:HG13	1.79	0.64
1:A:217[B]:PRO:O	1:A:218[B]:HIS:CB	2.46	0.63
1:B:216[A]:MET:HB2	1:B:219[A]:GLN:HG3	1.82	0.62
1:B:267:LEU:HB3	1:B:268:PRO:HA	1.81	0.62
1:A:217[B]:PRO:O	1:A:218[B]:HIS:HB3	2.01	0.60
1:A:266:LEU:HD13	1:A:406:VAL:HG21	1.84	0.58
1:A:217[B]:PRO:O	1:A:218[B]:HIS:CD2	2.58	0.56
1:A:216[B]:MET:CB	1:A:217[B]:PRO:CD	2.63	0.55
1:A:218[B]:HIS:O	1:A:218[B]:HIS:ND1	2.40	0.55
1:A:68:ASN:ND2	1:B:314[B]:GLN:OE1	2.40	0.52
1:B:272:VAL:HG13	1:B:406:VAL:HG12	1.94	0.50
1:A:406:VAL:HG23	1:A:407:ILE:HG13	1.94	0.50
1:B:208:VAL:HG13	1:B:248:ILE:CD1	2.41	0.49
1:A:217[B]:PRO:O	1:A:218[B]:HIS:CG	2.66	0.48
1:A:183:PHE:O	1:B:396:ASN:HB2	2.14	0.48
1:B:208:VAL:HG22	1:B:248:ILE:HD12	1.98	0.45
1:A:208:VAL:HG13	1:A:248:ILE:CD1	2.48	0.43
1:A:400:TRP:CZ2	1:B:172:GLN:HB3	2.53	0.43
3:A:502:MLI:H11	1:B:305:GLN:HB2	2.00	0.43
1:A:104:PHE:CE2	1:A:171:PHE:HB2	2.53	0.43
1:A:305:GLN:CB	3:B:502:MLI:H12	2.48	0.42
1:A:144:ASN:HA	1:A:145:PRO:HA	1.90	0.42
1:B:65:TYR:CZ	1:B:82:LEU:HD11	2.54	0.42
1:A:305:GLN:HB2	3:B:502:MLI:H12	2.01	0.42
1:B:272:VAL:HG13	1:B:406:VAL:CG1	2.49	0.41
1:B:70:LYS:HE3	3:B:502:MLI:H11	2.03	0.41
1:B:352:VAL:O	1:B:360:ARG:NH2	2.54	0.40
1:B:78:GLN:HE22	3:B:502:MLI:C1	2.35	0.40

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	405/410 (99%)	385 (95%)	18 (4%)	2 (0%)	29	17
1	B	407/410 (99%)	389 (96%)	17 (4%)	1 (0%)	47	38
All	All	812/820 (99%)	774 (95%)	35 (4%)	3 (0%)	47	22

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	216[A]	MET
1	A	216[B]	MET
1	B	268	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	337/337 (100%)	336 (100%)	1 (0%)	92	92
1	B	338/337 (100%)	335 (99%)	3 (1%)	78	77
All	All	675/674 (100%)	671 (99%)	4 (1%)	88	85

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	71	SER
1	B	216[A]	MET
1	B	216[B]	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	316	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

4 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	501	-	21,29,29	1.06	1 (4%)	18,42,42	1.76	4 (22%)
3	MLI	B	502	-	0,6,6	0.00	-	0,7,7	0.00	-
2	SAM	A	501	-	21,29,29	0.99	0	18,42,42	1.92	3 (16%)
3	MLI	A	502	-	0,6,6	0.00	-	0,7,7	0.00	-

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	501	-	-	0/8/33/33	0/3/3/3
3	MLI	B	502	-	-	0/0/4/4	-
2	SAM	A	501	-	-	1/8/33/33	0/3/3/3
3	MLI	A	502	-	-	0/0/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	501	SAM	C5-C4	2.53	1.47	1.40

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	SAM	C5'-SD-CG	5.10	116.41	103.40
2	B	501	SAM	C5'-SD-CG	4.60	115.14	103.40
2	A	501	SAM	N3-C2-N1	-3.93	122.53	128.68
2	B	501	SAM	N3-C2-N1	-3.67	122.95	128.68
2	A	501	SAM	C2-N1-C6	2.74	123.44	118.75
2	B	501	SAM	C2-N1-C6	2.36	122.80	118.75
2	B	501	SAM	C1'-N9-C4	-2.21	122.76	126.64

There are no chirality outliers.

All (1) torsion outliers are listed below:
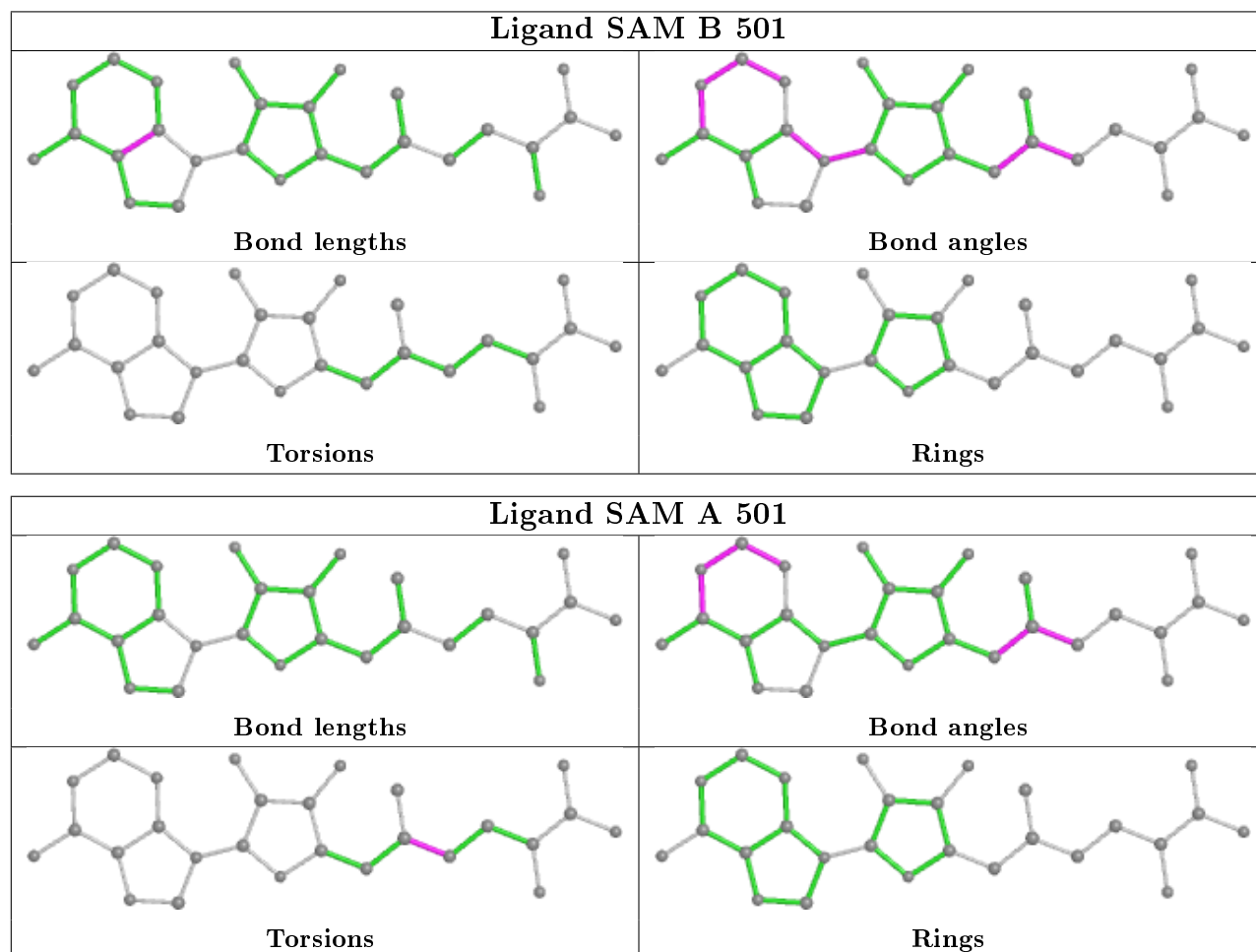
Mol	Chain	Res	Type	Atoms
2	A	501	SAM	CB-CG-SD-C5'

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	502	MLI	4	0
3	A	502	MLI	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 then 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.

Ligand SAM B 501 — Bond lengths, Bond angles, Torsions, Rings

Ligand SAM A 501 — Bond lengths, Bond angles, Torsions, Rings

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	401/410 (97%)	0.58	47 (11%)	4	7	24, 41, 85, 117	0
1	B	404/410 (98%)	0.60	45 (11%)	5	8	25, 42, 84, 120	0
All	All	805/820 (98%)	0.59	92 (11%)	5	8	24, 41, 86, 120	0

All (92) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	269	ALA	14.5
1	A	268	PRO	9.2
1	A	272	VAL	8.8
1	B	390	GLY	8.2
1	B	388	GLU	8.0
1	B	410	ILE	7.8
1	A	388	GLU	7.6
1	B	393	ALA	7.1
1	B	409	VAL	7.1
1	A	389	SER	7.1
1	B	389	SER	7.1
1	B	272	VAL	7.0
1	A	7	THR	6.0
1	B	391	GLU	6.0
1	B	7	THR	5.9
1	A	391	GLU	5.8
1	B	268	PRO	5.7
1	A	390	GLY	5.6
1	B	394	SER	5.5
1	B	392	GLU	5.2
1	A	386	ALA	5.2
1	B	269	ALA	5.2
1	A	267	LEU	5.2
1	B	270	GLY	4.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	394	SER	4.7
1	A	277	ALA	4.5
1	B	277	ALA	4.5
1	B	407	ILE	4.5
1	B	387	ARG	4.3
1	B	289	VAL	4.2
1	A	393	ALA	4.2
1	A	270	GLY	4.2
1	B	406	VAL	4.2
1	A	407	ILE	4.1
1	A	289	VAL	4.0
1	A	406	VAL	3.9
1	B	386	ALA	3.9
1	A	392	GLU	3.9
1	A	343	ALA	3.8
1	B	179	ALA	3.8
1	B	187	ASP	3.7
1	B	308	ASP	3.7
1	B	69	GLY	3.7
1	A	178	ILE	3.7
1	A	385	ALA	3.6
1	B	385	ALA	3.6
1	B	288	ASP	3.5
1	A	179	ALA	3.4
1	A	276	LYS	3.4
1	A	308	ASP	3.4
1	B	401	VAL	3.3
1	B	384	ASP	3.3
1	B	276	LYS	3.3
1	B	273	PRO	3.2
1	B	178	ILE	3.2
1	B	217[A]	PRO	3.2
1	A	405	GLY	3.2
1	A	288	ASP	3.1
1	A	387	ARG	3.1
1	A	273	PRO	3.0
1	A	404	VAL	3.0
1	A	187	ASP	3.0
1	A	396	ASN	3.0
1	B	408	GLY	2.9
1	B	396	ASN	2.9
1	A	384	ASP	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	44	ILE	2.9
1	B	271	GLN	2.9
1	B	369	TRP	2.8
1	A	346	ARG	2.8
1	A	69	GLY	2.8
1	B	184	THR	2.7
1	A	290	PRO	2.6
1	A	186	PHE	2.6
1	A	43	VAL	2.6
1	A	403	VAL	2.6
1	A	185	GLY	2.5
1	B	186	PHE	2.5
1	B	404	VAL	2.5
1	A	291	GLU	2.5
1	A	401	VAL	2.4
1	A	217[A]	PRO	2.4
1	B	290	PRO	2.4
1	A	44	ILE	2.4
1	B	291	GLU	2.3
1	B	405	GLY	2.1
1	A	395	GLN	2.1
1	A	234	PRO	2.1
1	A	184	THR	2.1
1	B	313	GLU	2.1
1	A	274	ASP	2.1
1	B	383	LEU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains (i)

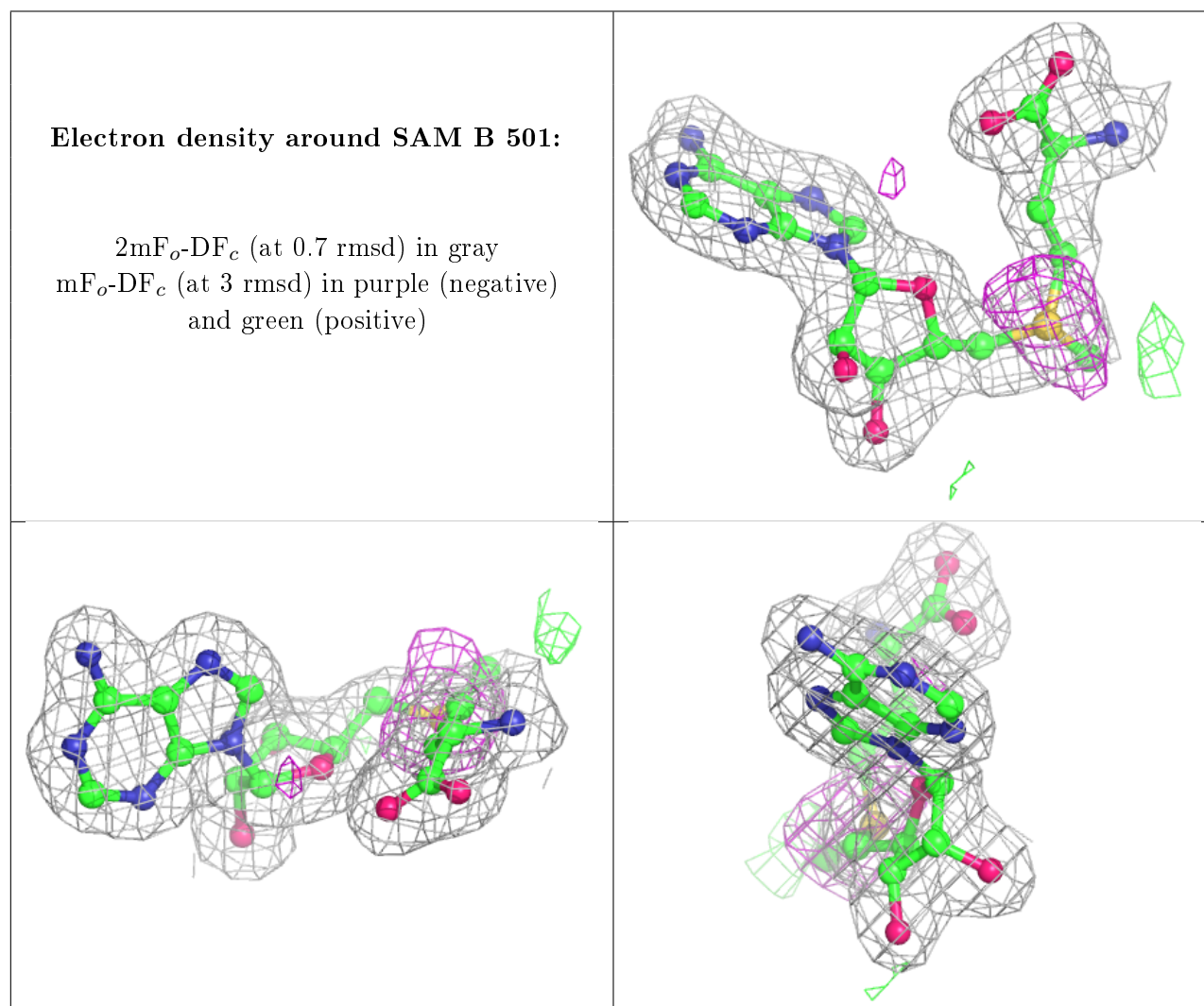There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

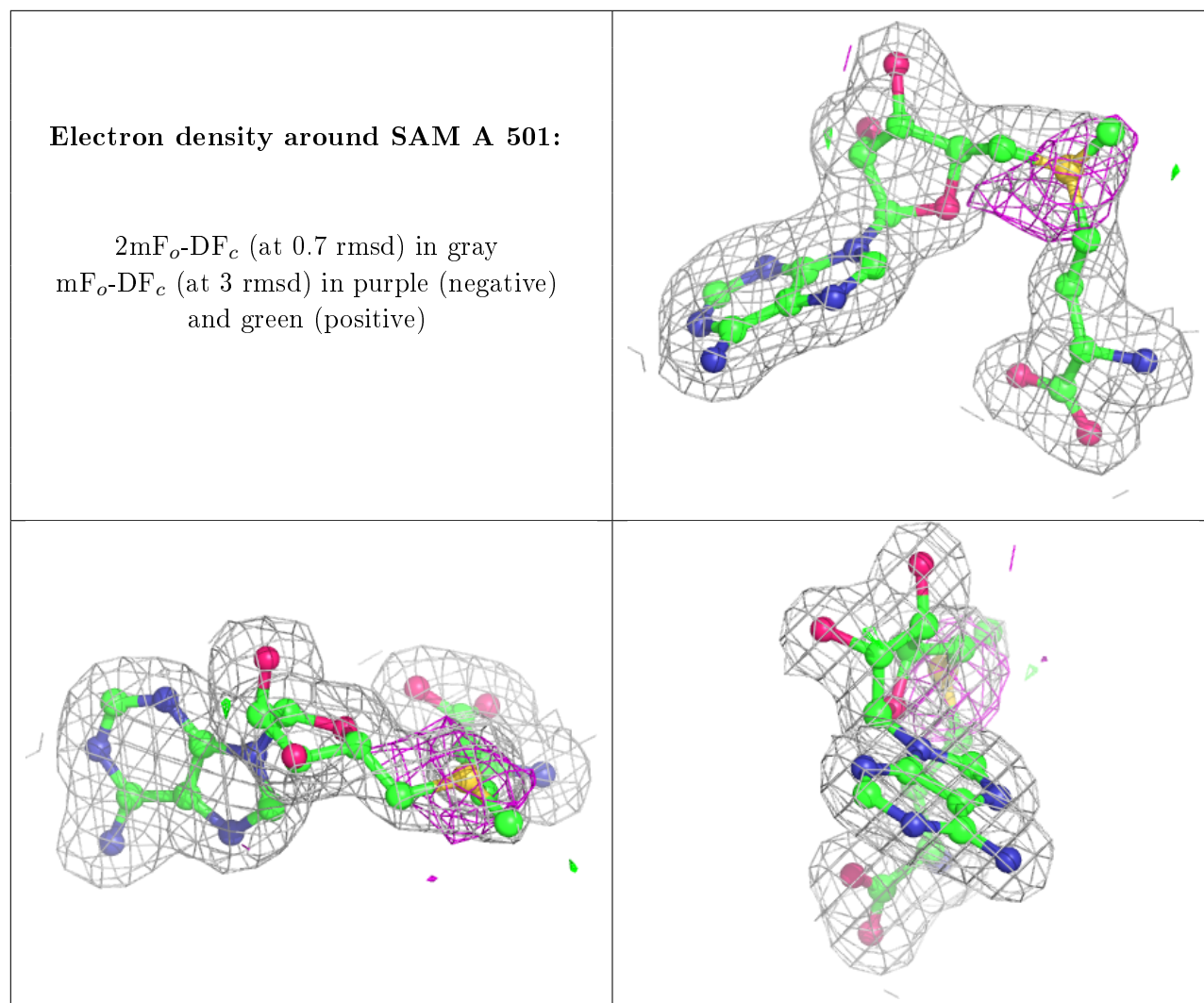## 6.4 Ligands (i)

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(Å^2)	Q<0.9
3	MLI	B	502	7/7	0.63	0.37	48,63,66,70	0
3	MLI	A	502	7/7	0.65	0.35	51,61,64,70	0
2	SAM	B	501	27/27	0.95	0.12	27,30,36,40	0
2	SAM	A	501	27/27	0.96	0.11	27,29,34,37	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.

Electron density around SAM B 501:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative) and green (positive)

Electron density around SAM A 501:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers ⓘ

There are no such residues in this entry.