



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 27, 2026 – 07:11 PM UTC

PDB ID : 4PIO / pdb_00004pio
Title : Ergothioneine-biosynthetic methyltransferase EgtD in complex with N,N-dimethylhistidine and SAH
Authors : Vit, A.; Seebeck, F.P.; Blankenfeldt, W.
Deposited on : 2014-05-09
Resolution : 1.51 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 ⓘ 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	:	FAILED
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

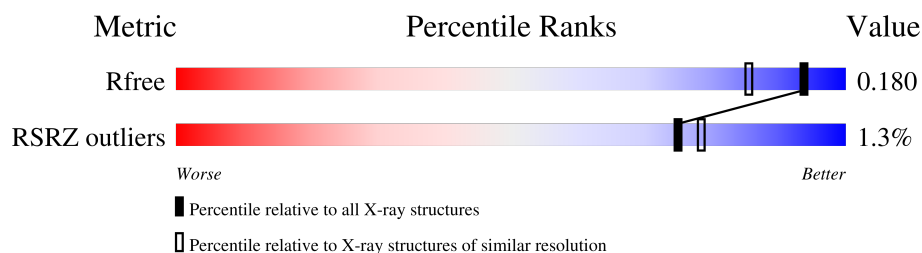
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.51 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	4037 (1.50-1.50)
RSRZ outliers	180081	4039 (1.50-1.50)

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10944 atoms, of which 4947 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 Histidine-specific methyltransferase EgtD.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	308	Total	C	H	N	O	S	0	10	1
			4842	1528	2414	429	465	6			
1	B	318	Total	C	H	N	O	S	0	8	0
			4969	1569	2470	444	480	6			

There are 12 discrepancies between the modelled and reference sequences:

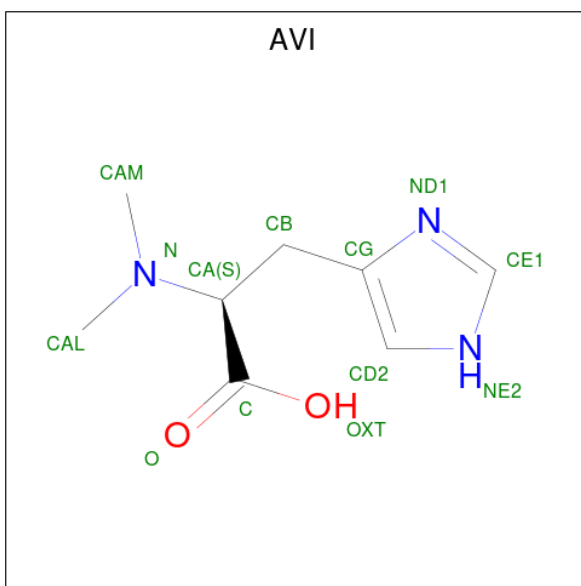
Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP A0R5M8
A	0	HIS	-	expression tag	UNP A0R5M8
A	2	ALA	THR	engineered mutation	UNP A0R5M8
A	29	THR	ALA	engineered mutation	UNP A0R5M8
A	30	GLN	PRO	engineered mutation	UNP A0R5M8
A	75	SER	ALA	engineered mutation	UNP A0R5M8
B	-1	GLY	-	expression tag	UNP A0R5M8
B	0	HIS	-	expression tag	UNP A0R5M8
B	2	ALA	THR	engineered mutation	UNP A0R5M8
B	29	THR	ALA	engineered mutation	UNP A0R5M8
B	30	GLN	PRO	engineered mutation	UNP A0R5M8
B	75	SER	ALA	engineered mutation	UNP A0R5M8

- Molecule 2 is S-ADENOSYL-L-HOMOCYSTEINE (CCD ID: SAH) (formula: $C_{14}H_{20}N_6O_5S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total 46	C 14	H 20	N 6	O 5	S 1	0	0
2	B	1	Total 45	C 14	H 19	N 6	O 5	S 1	0	0

- Molecule 3 is N,N-dimethyl-L-histidine (CCD ID: AVI) (formula: $\text{C}_8\text{H}_{13}\text{N}_3\text{O}_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 25	C 8	H 12	N 3	O 2	0	0
3	B	1	Total 25	C 8	H 12	N 3	O 2	0	0

- Molecule 4 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	4	Total 4	Mg 4	0	0
4	B	2	Total 2	Mg 2	0	0

- Molecule 5 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total 1	Cl 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	475	Total 475	O 475	0	0
6	B	510	Total 510	O 510	0	0

MolProbity failed to run properly - this section is therefore empty.

3 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	56.71Å 67.51Å 79.68Å 90.00° 110.98° 90.00°	Depositor
Resolution (Å)	49.99 – 1.51 49.99 – 1.51	Depositor EDS
% Data completeness (in resolution range)	98.7 (49.99-1.51) 98.6 (49.99-1.51)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.09 (at 1.50Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.148 , 0.180 0.149 , 0.180	Depositor DCC
R_{free} test set	4356 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å ²)	10.8	Xtriage
Anisotropy	0.100	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 43.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.023 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	10944	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.52% 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.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

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

2 non-standard protein/DNA/RNA residues 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	$\# Z > 2$
1	CSD	B	285	1	4,7,8	6.25	1 (25%)	1,8,10	0.99	0
1	CSD	A	285	1	4,7,8	5.53	1 (25%)	1,8,10	0.69	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
1	CSD	B	285	1	-	1/2/6/8	-
1	CSD	A	285	1	-	1/2/6/8	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	285	CSD	OD1-SG	-12.44	1.36	1.47
1	A	285	CSD	OD1-SG	-10.84	1.37	1.47

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	285	CSD	CA-CB-SG-OD1
1	B	285	CSD	CA-CB-SG-OD1

There are no ring outliers.

No monomer is involved in short contacts.

4.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

4.6 Ligand geometry [i](#)

Of 11 ligands modelled in this entry, 7 are monoatomic - leaving 4 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
2	SAH	A	401	-	27,28,28	0.89	1 (3%)	36,40,40	1.68	6 (16%)
3	AVI	A	402	-	12,13,13	1.33	2 (16%)	14,17,17	1.32	3 (21%)
3	AVI	B	402	-	12,13,13	1.34	2 (16%)	14,17,17	0.82	0
2	SAH	B	401	-	27,28,28	0.94	2 (7%)	36,40,40	1.74	8 (22%)

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	401	-	-	0/15/31/31	0/3/3/3
3	AVI	A	402	-	-	6/12/12/12	0/1/1/1
3	AVI	B	402	-	-	5/12/12/12	0/1/1/1
2	SAH	B	401	-	-	0/15/31/31	0/3/3/3

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	402	AVI	OXT-C	-3.92	1.18	1.30
3	B	402	AVI	OXT-C	-3.73	1.18	1.30
2	A	401	SAH	C5-N7	-2.42	1.34	1.39
3	B	402	AVI	CG-ND1	-2.34	1.32	1.38
3	A	402	AVI	CG-ND1	-2.09	1.33	1.38
2	B	401	SAH	C2-N1	2.06	1.37	1.33
2	B	401	SAH	C5-N7	-2.06	1.35	1.39

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	401	SAH	N3-C2-N1	-6.05	119.42	128.58
2	B	401	SAH	N3-C2-N1	-5.45	120.33	128.58
2	B	401	SAH	N9-C8-N7	-3.37	109.15	113.94
2	A	401	SAH	N9-C8-N7	-3.24	109.34	113.94
2	B	401	SAH	C5-C4-N3	-3.06	122.50	126.72
2	B	401	SAH	C5-N7-C8	3.03	108.21	103.45
2	A	401	SAH	C2-N1-C6	2.90	123.49	118.73
2	B	401	SAH	C2-N3-C4	2.79	118.65	111.83
2	B	401	SAH	C4-C5-N7	-2.78	107.41	110.58
2	A	401	SAH	C5-N7-C8	2.65	107.61	103.45
2	B	401	SAH	C5'-SD-CG	-2.53	94.74	102.26

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	401	SAH	C2-N3-C4	2.51	117.97	111.83
3	A	402	AVI	CAM-N-CAL	-2.43	103.36	110.49
2	B	401	SAH	C5-C4-N9	2.31	108.33	105.81
2	A	401	SAH	C5-C4-N3	-2.28	123.57	126.72
3	A	402	AVI	OXT-C-O	2.20	129.07	124.08
3	A	402	AVI	O-C-CA	-2.09	116.02	122.44

There are no chirality outliers.

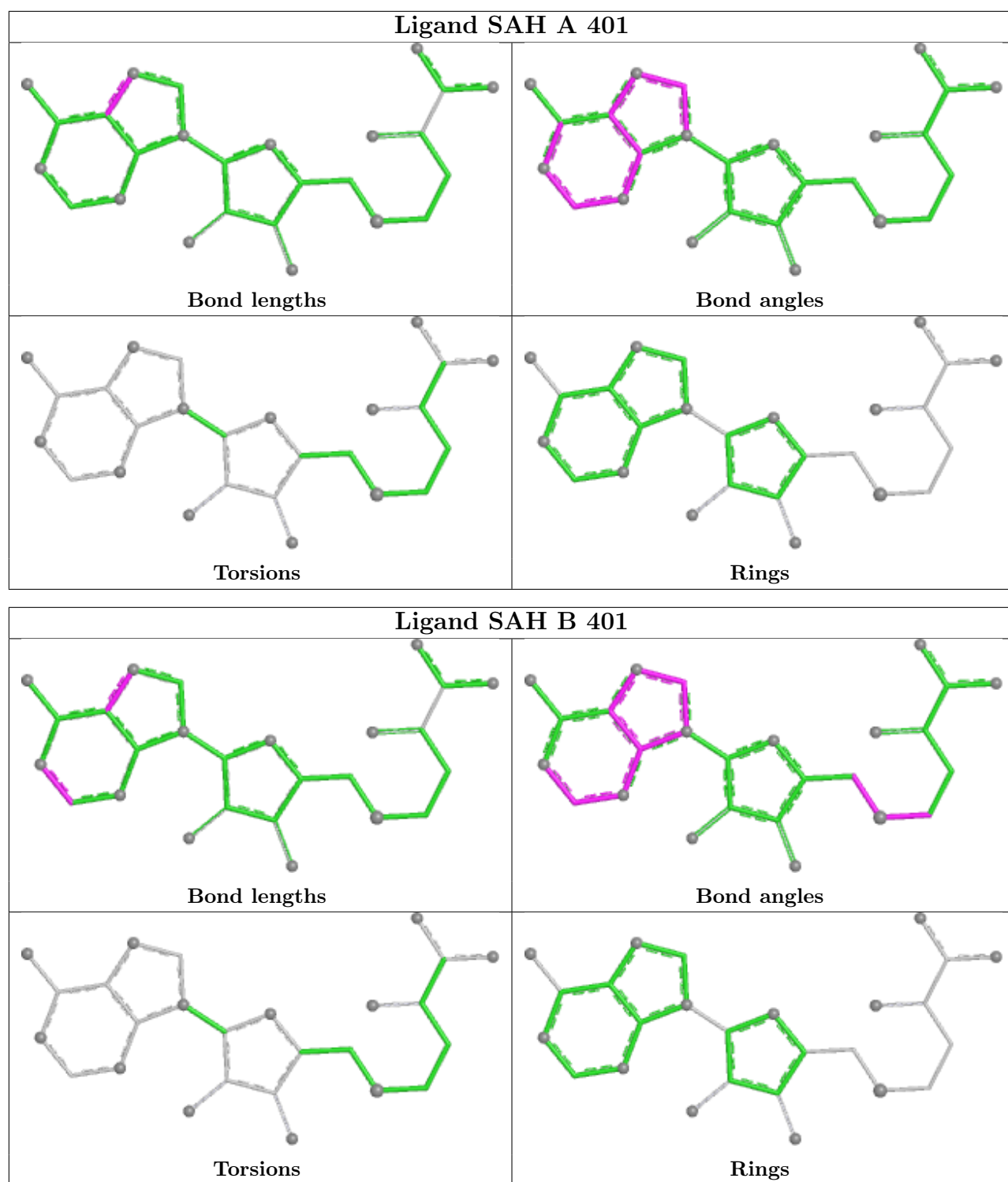
All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	402	AVI	C-CA-CB-CG
3	A	402	AVI	CA-CB-CG-CD2
3	B	402	AVI	CB-CA-N-CAM
3	A	402	AVI	CB-CA-N-CAM
3	B	402	AVI	O-C-CA-N
3	B	402	AVI	CA-CB-CG-CD2
3	A	402	AVI	CA-CB-CG-ND1
3	B	402	AVI	CA-CB-CG-ND1
3	A	402	AVI	N-CA-CB-CG
3	B	402	AVI	N-CA-CB-CG
3	A	402	AVI	O-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.



4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

5 Fit of model and data [i](#)

5.1 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	307/323 (95%)	-0.63	0	100 100	5, 12, 25, 32	11 (3%)
1	B	317/323 (98%)	-0.39	8 (2%)	58 62	4, 14, 30, 54	8 (2%)
All	All	624/646 (96%)	-0.51	8 (1%)	75 78	4, 13, 27, 54	19 (3%)

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	8	TYR	3.8
1	B	152	VAL	3.3
1	B	9	LEU	2.8
1	B	4	SER	2.8
1	B	6	ALA	2.7
1	B	151	HIS	2.7
1	B	11	ALA	2.3
1	B	10	ALA	2.3

5.2 Non-standard residues in protein, DNA, RNA chains [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
1	CSD	A	285	8/9	0.98	0.06	7,9,18,19	0
1	CSD	B	285	8/9	0.98	0.05	6,8,14,15	0

5.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.4 Ligands

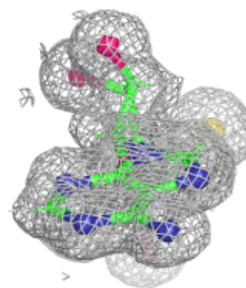
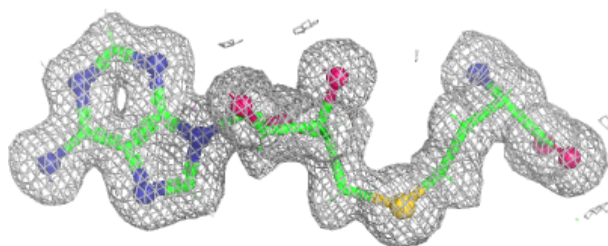
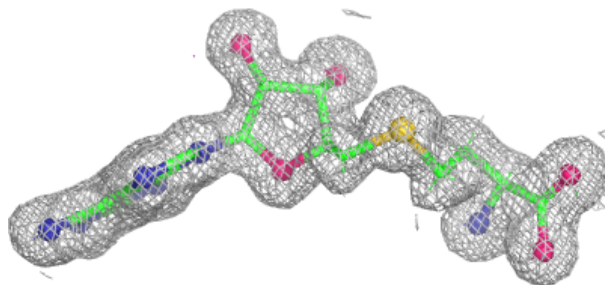
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	MG	A	406	1/1	0.90	0.08	40,40,40,40	0
4	MG	B	404	1/1	0.92	0.08	40,40,40,40	0
4	MG	A	403	1/1	0.94	0.07	33,33,33,33	0
5	CL	A	407	1/1	0.95	0.10	26,26,26,26	0
4	MG	A	404	1/1	0.96	0.11	36,36,36,36	0
3	AVI	B	402	13/13	0.97	0.04	5,7,11,11	0
3	AVI	A	402	13/13	0.97	0.04	6,7,9,9	0
4	MG	A	405	1/1	0.98	0.10	32,32,32,32	0
4	MG	B	403	1/1	0.99	0.08	19,19,19,19	0
2	SAH	A	401	26/26	0.99	0.03	4,6,10,11	0
2	SAH	B	401	26/26	0.99	0.04	5,8,10,11	0

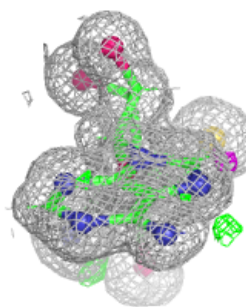
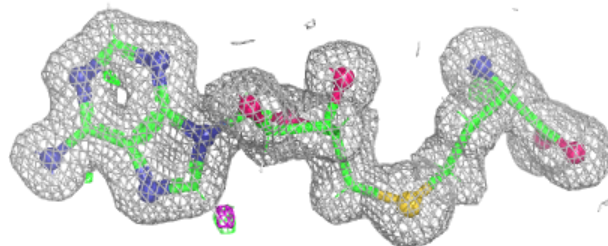
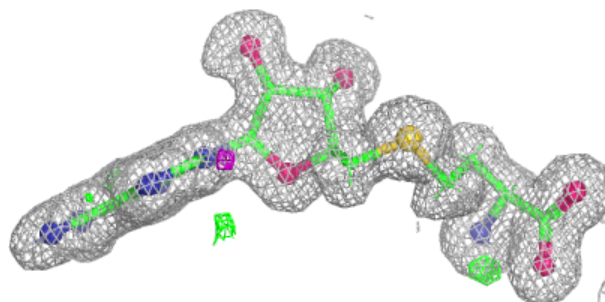
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around SAH A 401:

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

**Electron density around SAH B 401:**

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



5.5 Other polymers [i](#)

There are no such residues in this entry.