



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 16, 2024 – 12:39 AM EST

PDB ID : 6W8B  
Title : Structure of DNMT3A in complex with CGA DNA  
Authors : Anteneh, H.; Song, J.  
Deposited on : 2020-03-20  
Resolution : 2.40 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](mailto: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 : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtrriage (Phenix) : 1.21  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

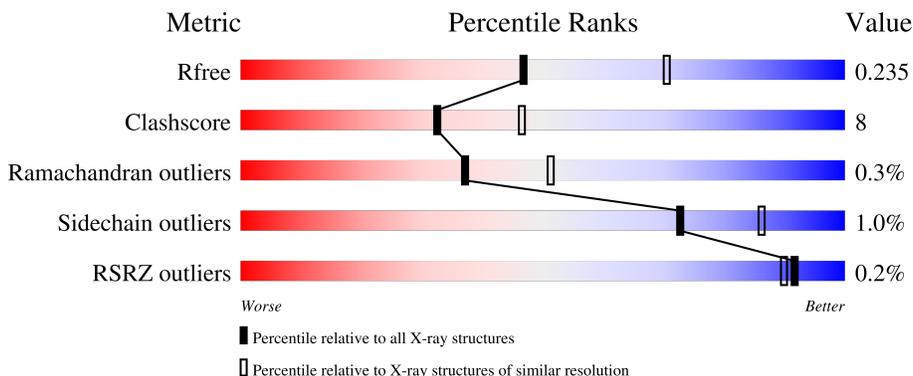
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	164625	4642 (2.40-2.40)
Clashscore	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)
RSRZ outliers	164620	4642 (2.40-2.40)

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  $\geq 3$ , 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	285	
1	D	285	
1	H	285	
1	K	285	
2	B	209	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	C	209	 67% 21% 13%
2	I	209	 67% 25% 7%
2	J	209	 66% 20% 14%
3	E	25	 68% 28% .
3	F	25	 60% 36% .
3	L	25	 64% 36%
3	M	25	 60% 36% .

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 17091 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 3A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	285	2277	1453	404	407	13	0	1	0
1	D	285	2299	1466	412	408	13	0	2	0
1	H	285	2292	1461	403	415	13	0	2	0
1	K	285	2289	1461	410	405	13	0	2	0

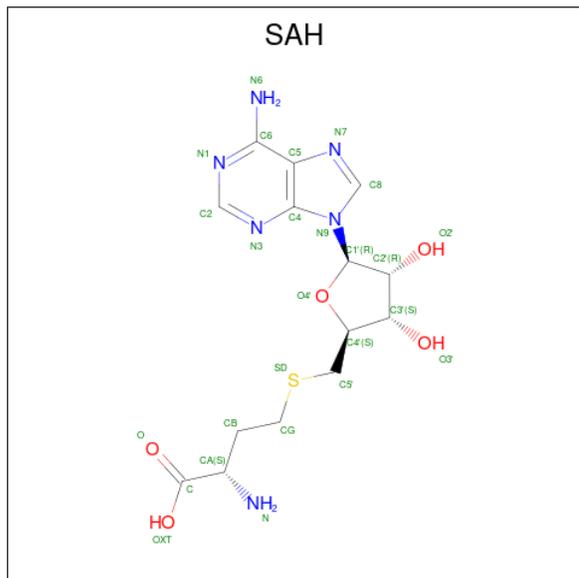
- Molecule 2 is a protein called DNA (cytosine-5)-methyltransferase 3-like.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	194	1474	964	240	266	4	0	0	0
2	C	182	1283	829	215	237	2	0	0	0
2	I	195	1480	968	244	265	3	0	0	0
2	J	180	1299	843	217	237	2	0	0	0

- Molecule 3 is a DNA chain called CGA DNA (25-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	E	25	493	235	88	146	24	0	0	0
3	F	25	510	245	93	148	24	0	0	0
3	L	25	493	235	88	146	24	0	0	0
3	M	25	510	245	93	148	24	0	0	0

- Molecule 4 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula:  $C_{14}H_{20}N_6O_5S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
4	A	1	Total 26	14	6	5	1	0	0
4	D	1	Total 26	14	6	5	1	0	0
4	H	1	Total 26	14	6	5	1	0	0
4	K	1	Total 26	14	6	5	1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	68	Total 68	68	0	0
5	B	6	Total 6	6	0	0
5	C	2	Total 2	2	0	0
5	D	65	Total 65	65	0	0
5	E	3	Total 3	3	0	0
5	F	3	Total 3	3	0	0

*Continued on next page...*

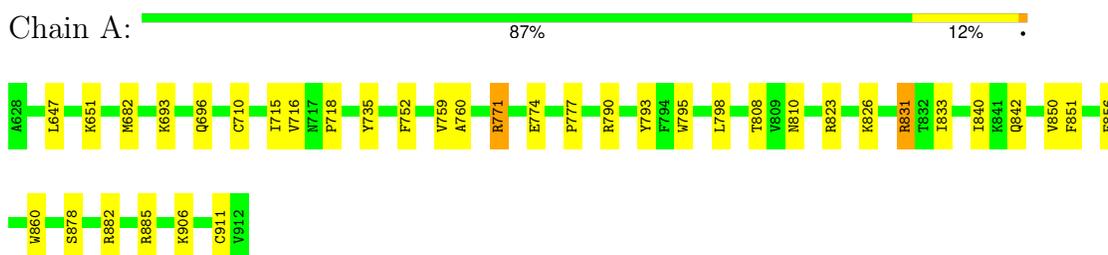
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
5	H	69	Total O 69 69	0	0
5	I	7	Total O 7 7	0	0
5	J	2	Total O 2 2	0	0
5	K	58	Total O 58 58	0	0
5	L	3	Total O 3 3	0	0
5	M	2	Total O 2 2	0	0

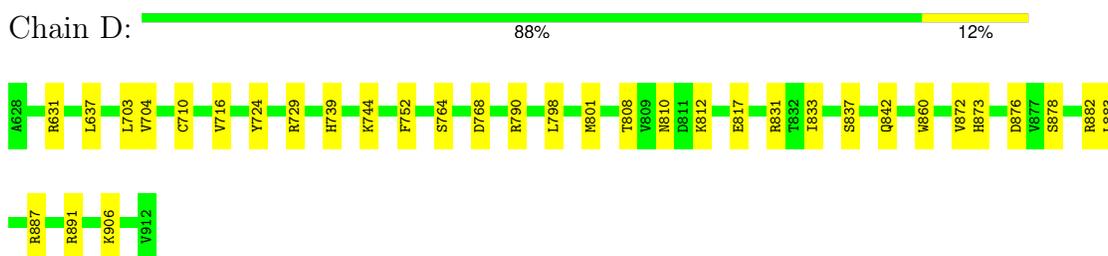
### 3 Residue-property plots [i](#)

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.

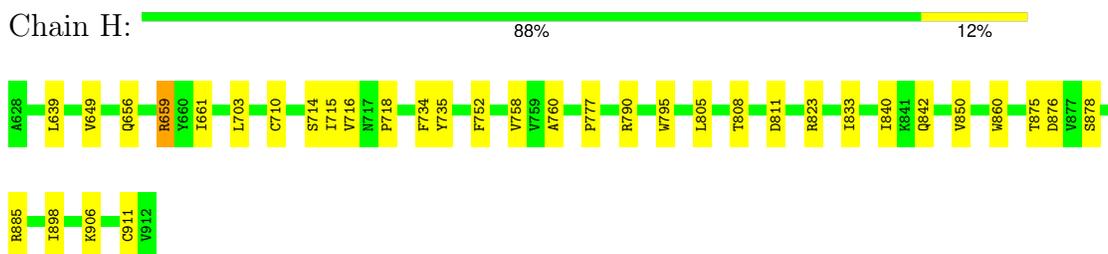
- Molecule 1: DNA (cytosine-5)-methyltransferase 3A



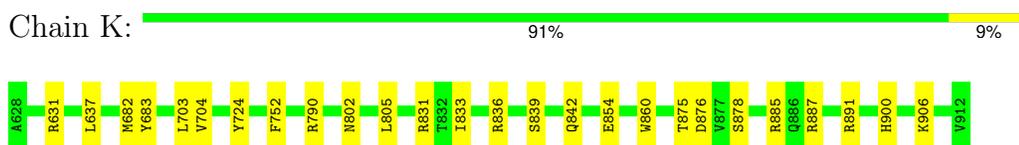
- Molecule 1: DNA (cytosine-5)-methyltransferase 3A



- Molecule 1: DNA (cytosine-5)-methyltransferase 3A

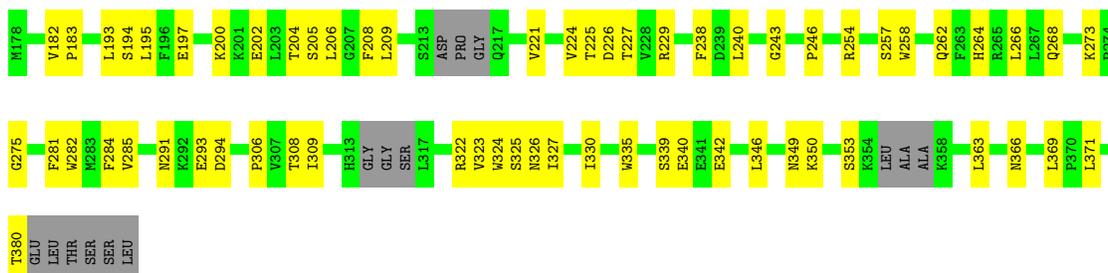


- Molecule 1: DNA (cytosine-5)-methyltransferase 3A



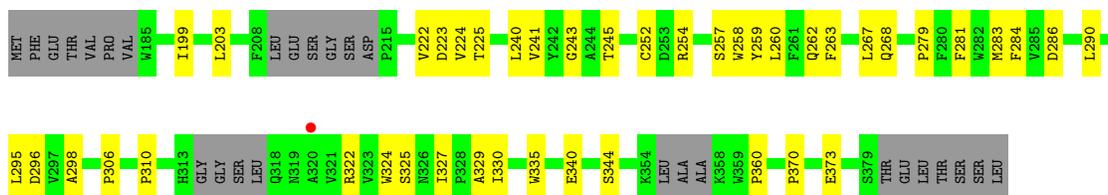
- Molecule 2: DNA (cytosine-5)-methyltransferase 3-like

Chain B:  63% 30% 7%



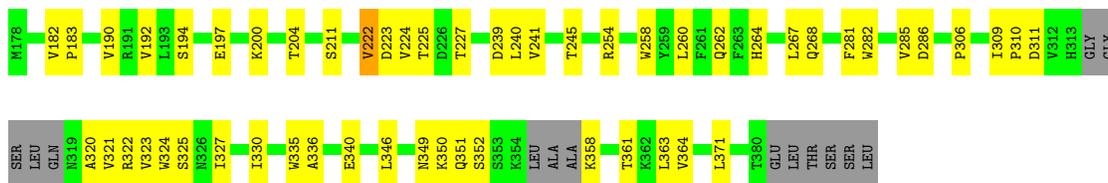
- Molecule 2: DNA (cytosine-5)-methyltransferase 3-like

Chain C:  67% 21% 13%



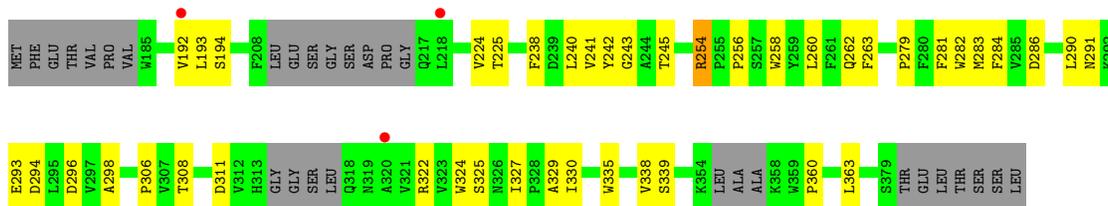
- Molecule 2: DNA (cytosine-5)-methyltransferase 3-like

Chain I:  67% 25% 7%



- Molecule 2: DNA (cytosine-5)-methyltransferase 3-like

Chain J:  66% 20% 14%



- Molecule 3: CGA DNA (25-MER)

Chain E:  68% 28% 4%



## ● Molecule 3: CGA DNA (25-MER)

Chain F:  60% 36%

## ● Molecule 3: CGA DNA (25-MER)

Chain L:  64% 36%

## ● Molecule 3: CGA DNA (25-MER)

Chain M:  60% 36%

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	186.35Å 186.35Å 81.67Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.87 – 2.40 48.87 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.87-2.40) 99.9 (48.87-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.28 (at 2.39Å)	Xtrriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, $R_{free}$	0.212 , 0.233 0.214 , 0.235	Depositor DCC
$R_{free}$ test set	121893 reflections (1.62%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.2	Xtrriage
Anisotropy	0.605	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 63.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.488 for -h,-k,l 0.022 for h,-h-k,-l 0.022 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	17091	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	95.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.05% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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, PYO

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	0.27	0/2331	0.45	1/3153 (0.0%)
1	D	0.26	0/2356	0.42	0/3182
1	H	0.27	0/2346	0.43	0/3173
1	K	0.27	0/2349	0.44	1/3174 (0.0%)
2	B	0.28	0/1520	0.49	0/2089
2	C	0.28	0/1320	0.46	0/1817
2	I	0.27	0/1529	0.45	0/2101
2	J	0.31	0/1338	0.51	0/1842
3	E	0.53	0/530	0.95	0/815
3	F	0.55	0/550	0.91	0/846
3	L	0.55	0/530	0.96	0/815
3	M	0.56	0/550	0.95	0/846
All	All	0.32	0/17249	0.55	2/23853 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	K	682	MET	CA-CB-CG	-5.37	104.17	113.30
1	A	771	ARG	NE-CZ-NH1	5.16	122.88	120.30

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	2277	0	2221	27	0
1	D	2299	0	2275	26	0
1	H	2292	0	2229	26	0
1	K	2289	0	2263	20	0
2	B	1474	0	1307	38	0
2	C	1283	0	1053	29	0
2	I	1480	0	1315	31	0
2	J	1299	0	1078	33	0
3	E	493	0	270	11	0
3	F	510	0	279	11	0
3	L	493	0	270	11	0
3	M	510	0	279	10	0
4	A	26	0	19	1	0
4	D	26	0	19	0	0
4	H	26	0	19	2	0
4	K	26	0	19	0	0
5	A	68	0	0	2	0
5	B	6	0	0	1	0
5	C	2	0	0	0	0
5	D	65	0	0	1	0
5	E	3	0	0	0	0
5	F	3	0	0	0	0
5	H	69	0	0	1	0
5	I	7	0	0	1	0
5	J	2	0	0	0	0
5	K	58	0	0	2	0
5	L	3	0	0	0	0
5	M	2	0	0	0	0
All	All	17091	0	14915	245	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 8.

The worst 5 of 245 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)
2:J:327:ILE:HD11	2:J:330:ILE:HG13	1.44	0.97
2:J:281:PHE:HD2	2:J:327:ILE:HG22	1.31	0.95
2:J:281:PHE:CD2	2:J:327:ILE:HG22	2.08	0.89

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:808:THR:HG22	1:D:810:ASN:H	1.39	0.88
2:C:327:ILE:HD11	2:C:330:ILE:HG13	1.58	0.85

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	284/285 (100%)	275 (97%)	9 (3%)	0	100	100
1	D	285/285 (100%)	277 (97%)	8 (3%)	0	100	100
1	H	285/285 (100%)	277 (97%)	8 (3%)	0	100	100
1	K	285/285 (100%)	276 (97%)	9 (3%)	0	100	100
2	B	186/209 (89%)	172 (92%)	14 (8%)	0	100	100
2	C	174/209 (83%)	158 (91%)	13 (8%)	3 (2%)	7	10
2	I	189/209 (90%)	174 (92%)	14 (7%)	1 (0%)	25	38
2	J	172/209 (82%)	159 (92%)	12 (7%)	1 (1%)	22	33
All	All	1860/1976 (94%)	1768 (95%)	87 (5%)	5 (0%)	37	51

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	360	PRO
2	J	360	PRO
2	I	222	VAL
2	C	310	PRO
2	C	370	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	239/250 (96%)	235 (98%)	4 (2%)	56	75
1	D	246/250 (98%)	245 (100%)	1 (0%)	89	95
1	H	242/250 (97%)	240 (99%)	2 (1%)	79	90
1	K	244/250 (98%)	244 (100%)	0	100	100
2	B	142/191 (74%)	140 (99%)	2 (1%)	62	79
2	C	107/191 (56%)	106 (99%)	1 (1%)	75	88
2	I	142/191 (74%)	140 (99%)	2 (1%)	62	79
2	J	111/191 (58%)	107 (96%)	4 (4%)	30	49
All	All	1473/1764 (84%)	1457 (99%)	16 (1%)	73	84

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	J	294	ASP
2	J	254	ARG
1	H	659	ARG
2	J	194	SER
1	D	764	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	H	656	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](#)

4 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
3	PYO	E	427	3	16,20,21	2.15	4 (25%)	21,28,31	1.07	1 (4%)
3	PYO	M	427	3	16,20,21	2.13	4 (25%)	21,28,31	1.09	2 (9%)
3	PYO	F	427	3	16,20,21	2.14	4 (25%)	21,28,31	1.02	1 (4%)
3	PYO	L	427	3	16,20,21	2.14	4 (25%)	21,28,31	1.00	2 (9%)

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	PYO	E	427	3	-	4/7/25/26	0/2/2/2
3	PYO	M	427	3	-	4/7/25/26	0/2/2/2
3	PYO	F	427	3	-	4/7/25/26	0/2/2/2
3	PYO	L	427	3	-	4/7/25/26	0/2/2/2

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	427	PYO	C6-C5	6.79	1.49	1.35
3	E	427	PYO	C6-C5	6.76	1.49	1.35
3	L	427	PYO	C6-C5	6.74	1.49	1.35
3	M	427	PYO	C6-C5	6.71	1.49	1.35
3	E	427	PYO	C2-N1	3.05	1.46	1.40

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	427	PYO	C5-C4-N3	-3.70	119.51	124.25

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	M	427	PYO	C5-C4-N3	-3.62	119.62	124.25
3	F	427	PYO	C5-C4-N3	-3.59	119.65	124.25
3	L	427	PYO	C5-C4-N3	-3.29	120.03	124.25
3	M	427	PYO	C5-C6-N1	-2.14	119.01	121.24

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	427	PYO	C2'-C1'-N1-C6
3	E	427	PYO	C2'-C1'-N1-C2
3	L	427	PYO	C2'-C1'-N1-C6
3	L	427	PYO	C2'-C1'-N1-C2
3	M	427	PYO	C2'-C1'-N1-C6

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	427	PYO	1	0
3	M	427	PYO	1	0
3	F	427	PYO	1	0

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides 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
4	SAH	K	1001	-	23,28,28	1.27	3 (13%)	22,40,40	1.80	4 (18%)
4	SAH	D	1001	-	23,28,28	1.28	3 (13%)	22,40,40	1.89	4 (18%)
4	SAH	H	1001	-	23,28,28	1.26	3 (13%)	22,40,40	1.98	4 (18%)
4	SAH	A	1001	-	23,28,28	1.25	3 (13%)	22,40,40	1.97	5 (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
4	SAH	K	1001	-	-	4/11/31/31	0/3/3/3
4	SAH	D	1001	-	-	4/11/31/31	0/3/3/3
4	SAH	H	1001	-	-	5/11/31/31	0/3/3/3
4	SAH	A	1001	-	-	5/11/31/31	0/3/3/3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1001	SAH	C2-N3	4.28	1.38	1.32
4	K	1001	SAH	C2-N3	4.23	1.38	1.32
4	H	1001	SAH	C2-N3	4.11	1.38	1.32
4	A	1001	SAH	C2-N3	4.04	1.38	1.32
4	D	1001	SAH	C2-N1	2.63	1.38	1.33

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	1001	SAH	N3-C2-N1	-6.48	119.88	128.67
4	A	1001	SAH	N3-C2-N1	-6.44	119.93	128.67
4	K	1001	SAH	N3-C2-N1	-6.36	120.05	128.67
4	D	1001	SAH	N3-C2-N1	-6.32	120.09	128.67
4	H	1001	SAH	C5'-SD-CG	-3.63	91.50	102.26

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1001	SAH	N-CA-CB-CG
4	D	1001	SAH	N-CA-CB-CG

*Continued on next page...*

*Continued from previous page...*

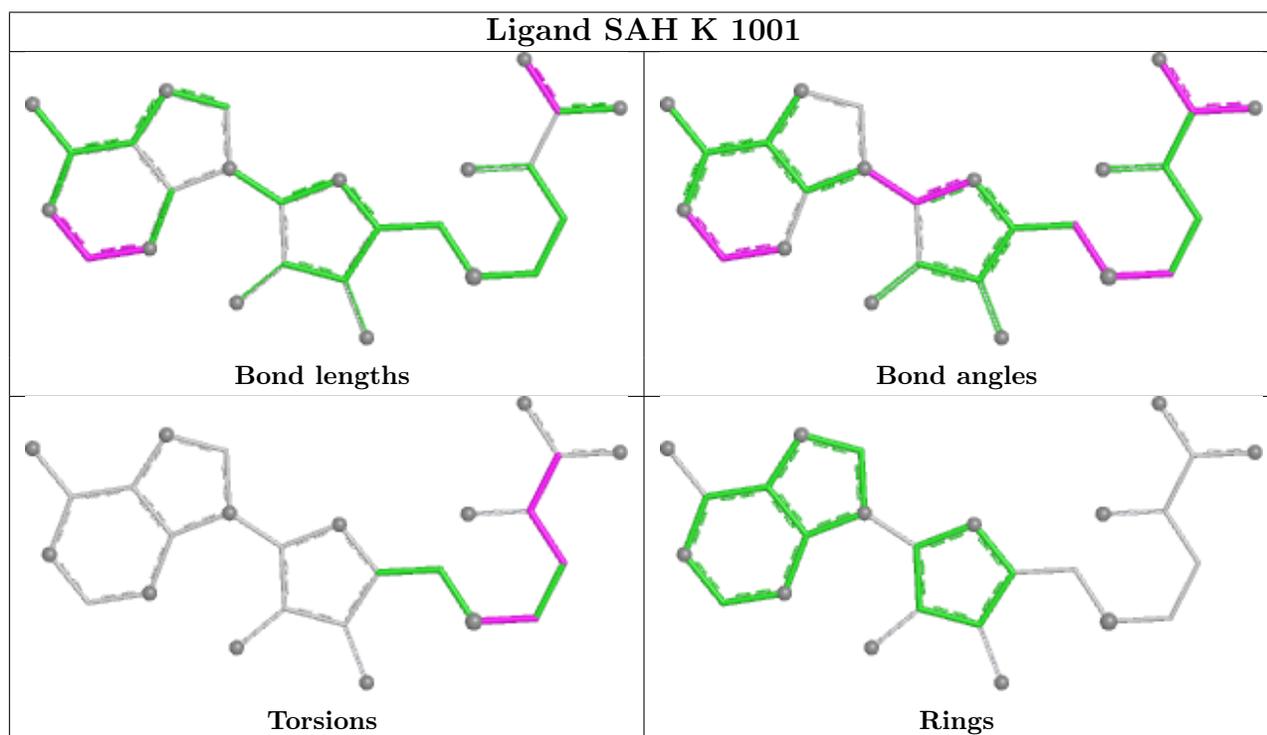
Mol	Chain	Res	Type	Atoms
4	H	1001	SAH	N-CA-CB-CG
4	K	1001	SAH	N-CA-CB-CG
4	D	1001	SAH	CB-CG-SD-C5'

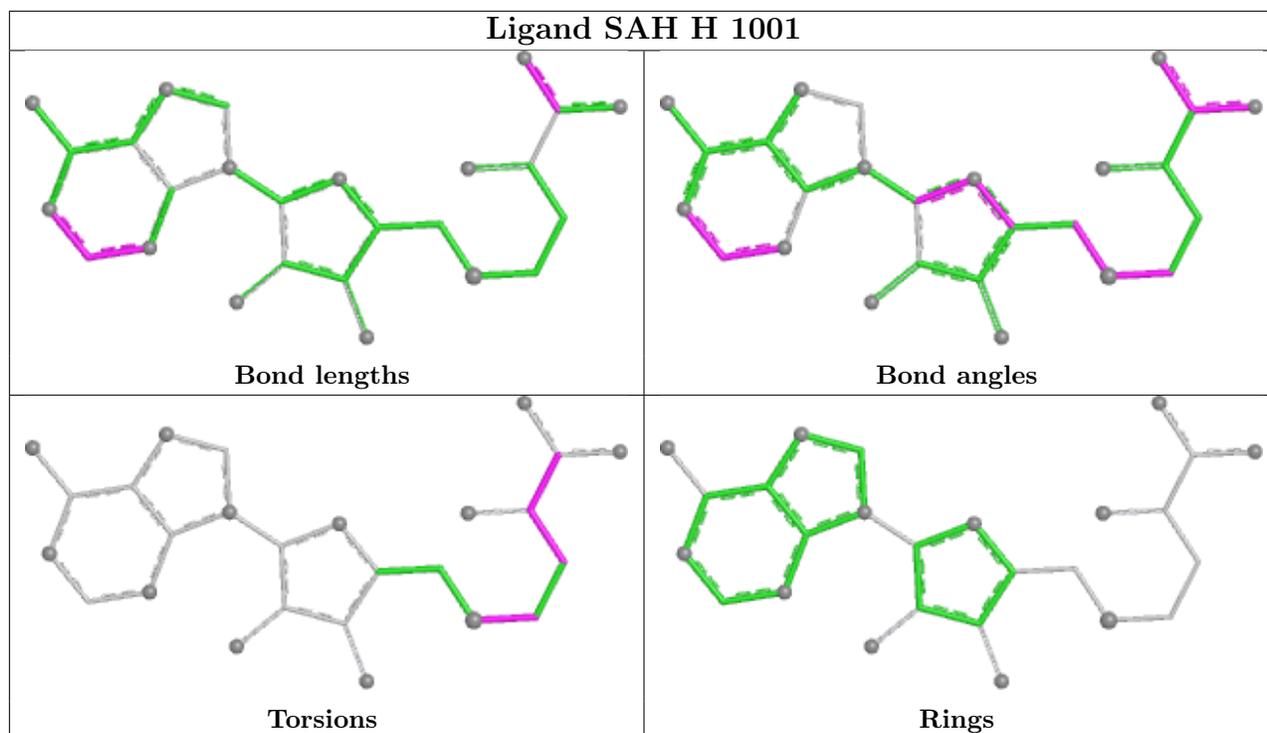
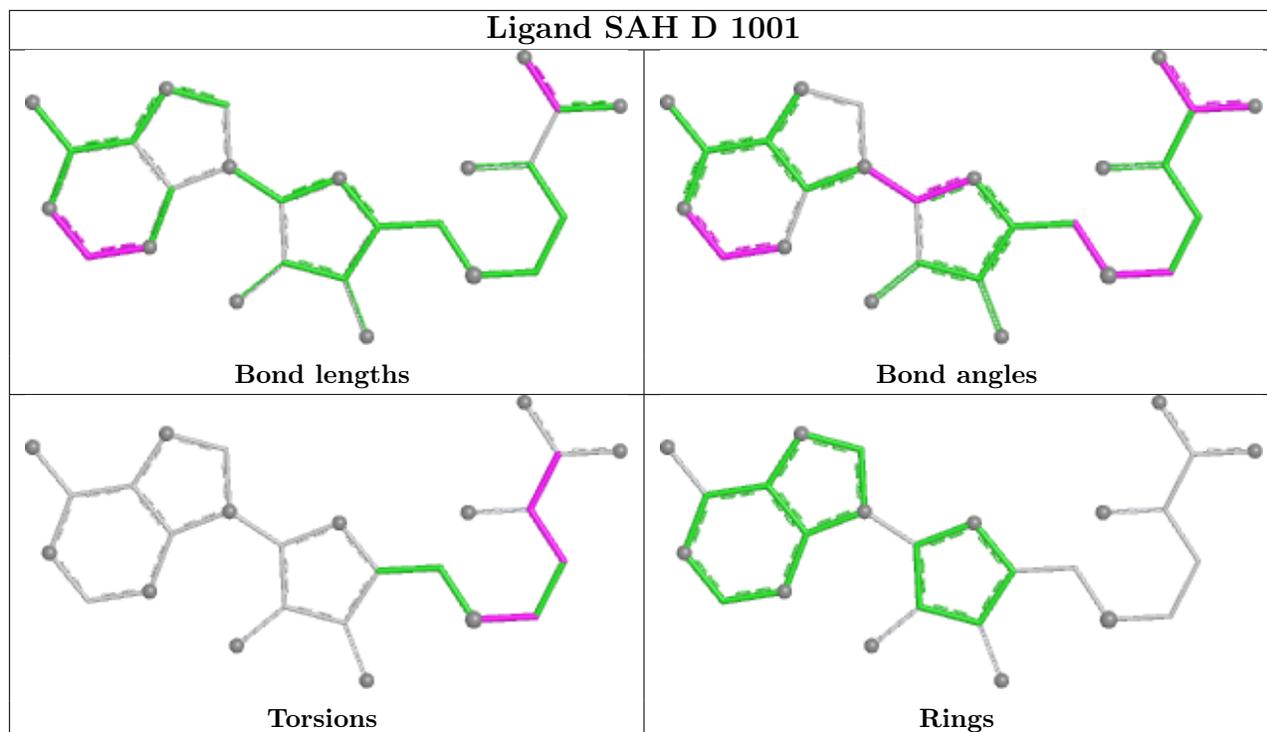
There are no ring outliers.

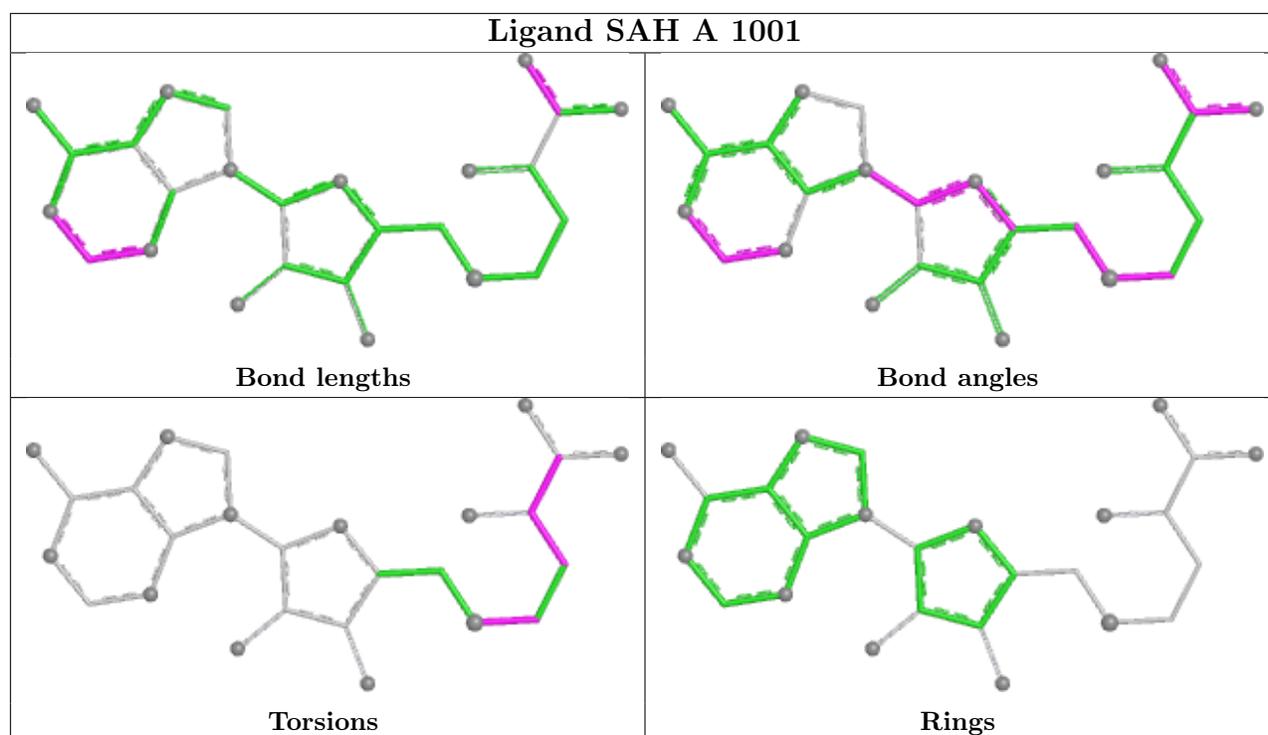
2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	1001	SAH	2	0
4	A	1001	SAH	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 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	285/285 (100%)	-1.39	0 100 100	37, 69, 99, 152	1 (0%)
1	D	285/285 (100%)	-1.41	0 100 100	34, 67, 99, 155	2 (0%)
1	H	285/285 (100%)	-1.40	0 100 100	34, 69, 98, 154	2 (0%)
1	K	285/285 (100%)	-1.39	0 100 100	36, 67, 95, 165	2 (0%)
2	B	194/209 (92%)	-0.59	0 100 100	69, 118, 194, 217	0
2	C	182/209 (87%)	-0.44	1 (0%) 87 85	78, 147, 203, 230	0
2	I	195/209 (93%)	-0.63	0 100 100	69, 118, 186, 211	0
2	J	180/209 (86%)	-0.43	3 (1%) 69 65	79, 142, 201, 218	0
3	E	24/25 (96%)	-1.35	0 100 100	74, 121, 146, 184	0
3	F	24/25 (96%)	-1.38	0 100 100	87, 113, 148, 174	0
3	L	24/25 (96%)	-1.37	0 100 100	73, 120, 145, 201	0
3	M	24/25 (96%)	-1.42	0 100 100	84, 114, 148, 184	0
All	All	1987/2076 (95%)	-1.07	4 (0%) 92 90	34, 84, 181, 230	7 (0%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	320	ALA	3.8
2	J	218	LEU	2.5
2	C	320	ALA	2.3
2	J	192	VAL	2.0

### 6.2 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	PYO	F	427	19/20	0.99	0.04	72,76,86,91	0
3	PYO	E	427	19/20	1.00	0.03	62,71,79,84	0
3	PYO	L	427	19/20	1.00	0.03	61,67,80,81	0
3	PYO	M	427	19/20	1.00	0.04	69,78,90,95	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

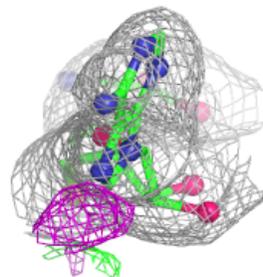
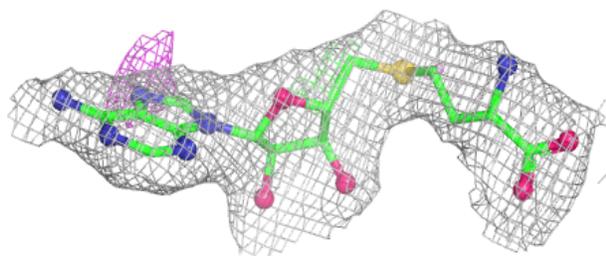
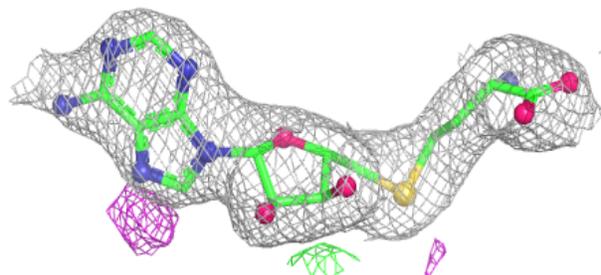
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SAH	A	1001	26/26	0.99	0.04	50,66,74,76	0
4	SAH	D	1001	26/26	0.99	0.04	53,68,98,119	0
4	SAH	H	1001	26/26	0.99	0.04	49,64,71,73	0
4	SAH	K	1001	26/26	1.00	0.03	48,64,74,78	0

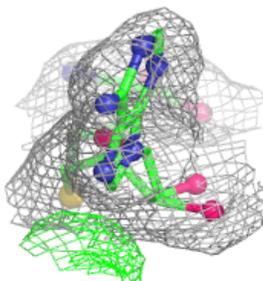
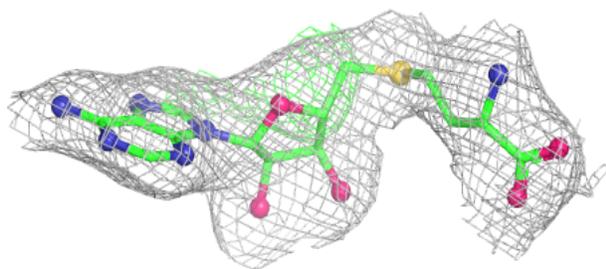
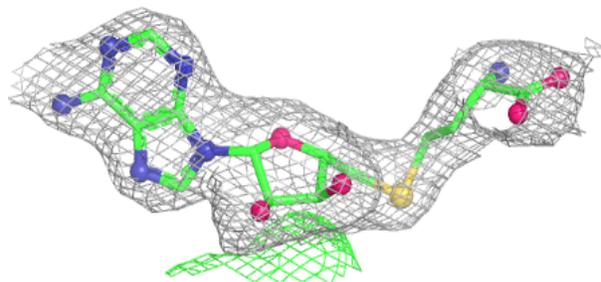
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around SAH A 1001:**

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

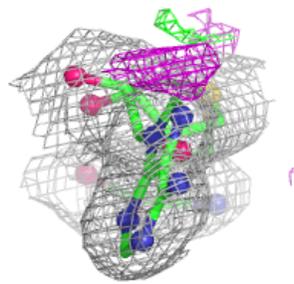
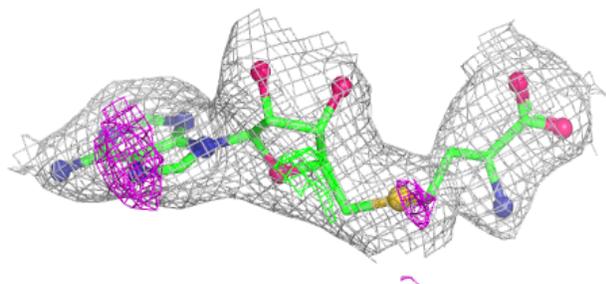
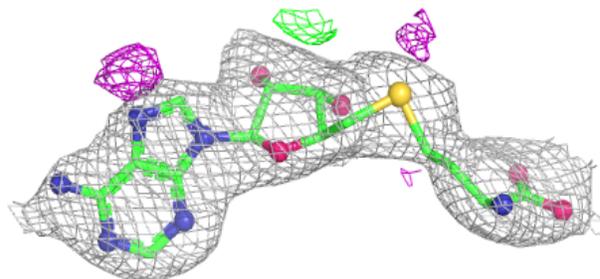
**Electron density around SAH D 1001:**

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

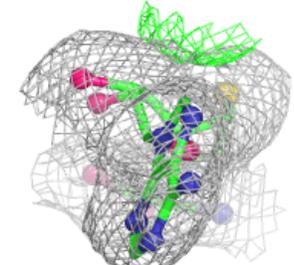
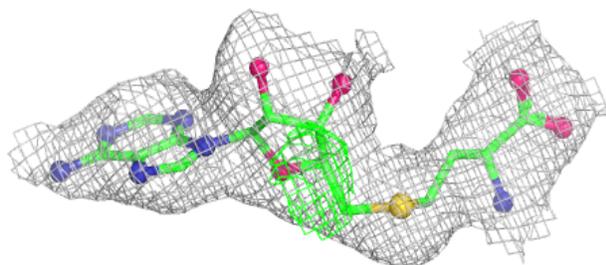
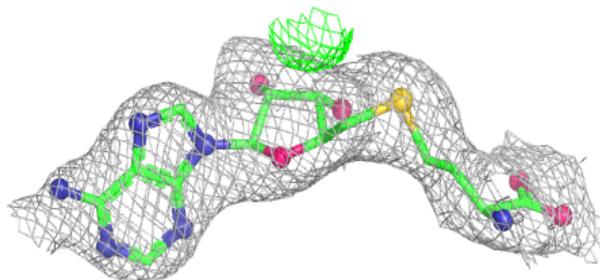


**Electron density around SAH H 1001:**

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

**Electron density around SAH K 1001:**

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



## 6.5 Other polymers [i](#)

There are no such residues in this entry.