



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 07:07 am BST

PDB ID : 5YX2  
Title : Crystal structure of DNMT3A-DNMT3L in complex with DNA containing two CpG sites  
Authors : Zhang, Z.M.; Song, J.  
Deposited on : 2017-12-01  
Resolution : 2.65 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](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 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

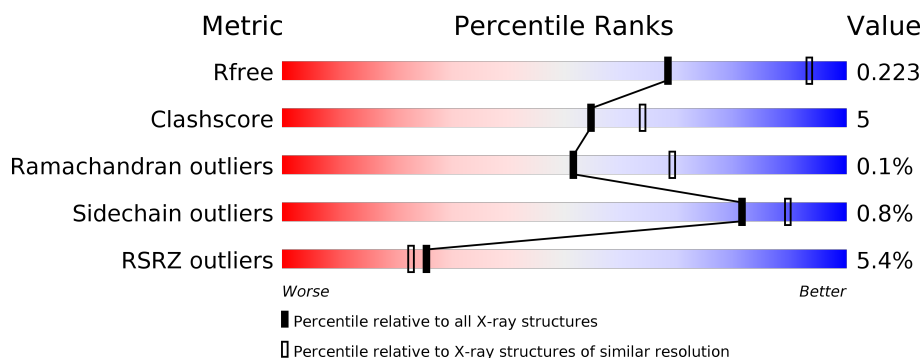
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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  $\geq 3$ , 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	<div> <div>2%</div> <div>84%</div> <div>16%</div> </div>
1	D	285	<div> <div>%</div> <div>91%</div> <div>9%</div> </div>
2	B	208	<div> <div>10%</div> <div>80%</div> <div>10%</div> <div>10%</div> </div>
2	C	208	<div> <div>12%</div> <div>70%</div> <div>14%</div> <div>16%</div> </div>
3	E	25	<div> <div>60%</div> <div>40%</div> </div>
3	F	25	<div> <div>88%</div> <div>12%</div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8380 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
1	A	285	Total	C	N	O	S	0	0	0
			2267	1450	402	402	13			
1	D	285	Total	C	N	O	S	0	0	0
			2269	1449	402	405	13			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	188	Total	C	N	O	S	0	0	0
			1398	917	233	245	3			
2	C	174	Total	C	N	O	S	0	0	0
			1263	821	211	229	2			

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

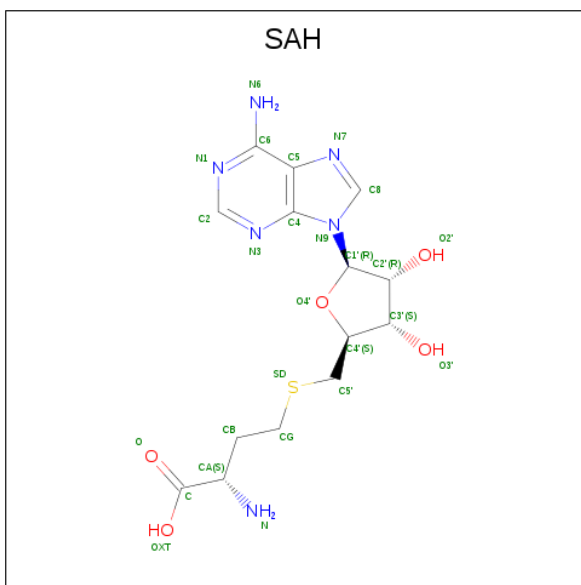
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	25	Total	C	N	O	P	0	0	0
			510	245	93	148	24			
3	F	25	Total	C	N	O	P	0	0	0
			510	245	93	148	24			

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula:  $C_{14}H_{20}N_6O_5S$ ) (labeled as "Ligand of Interest" by author).



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

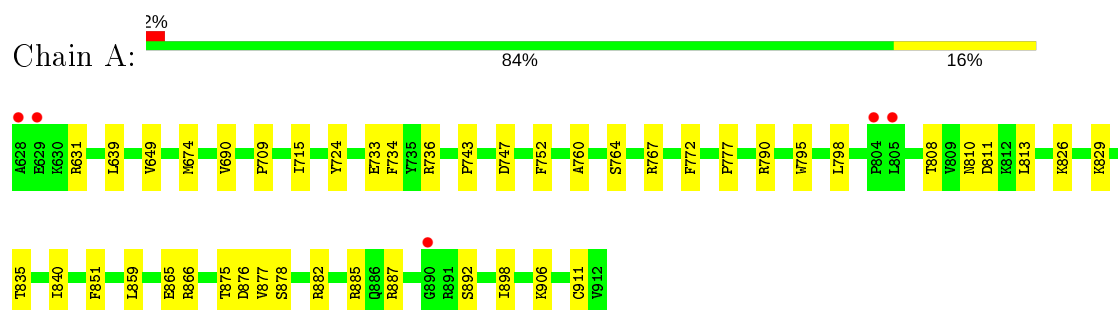
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	22	Total O 22 22	0	0
6	B	5	Total O 5 5	0	0
6	D	38	Total O 38 38	0	0
6	C	2	Total O 2 2	0	0
6	E	1	Total O 1 1	0	0
6	F	7	Total O 7 7	0	0

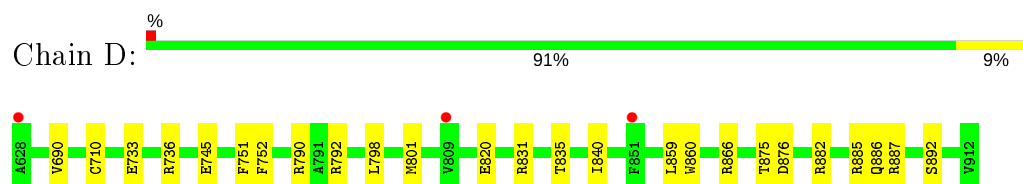
### 3 Residue-property plots [i](#)

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.

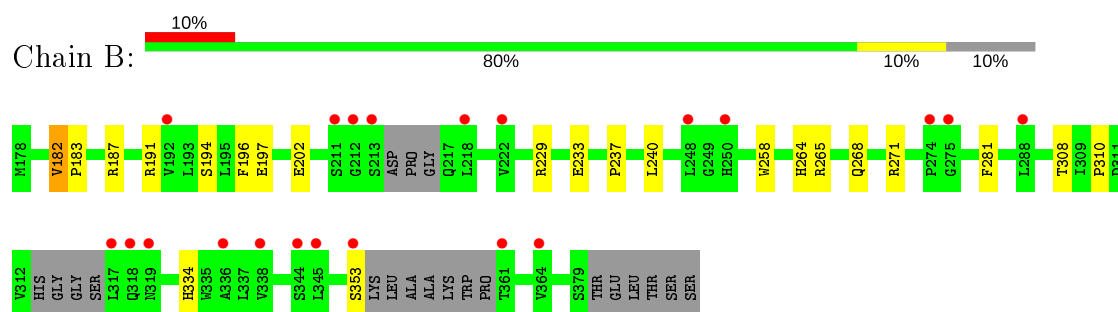
- Molecule 1: DNA (cytosine-5)-methyltransferase 3A



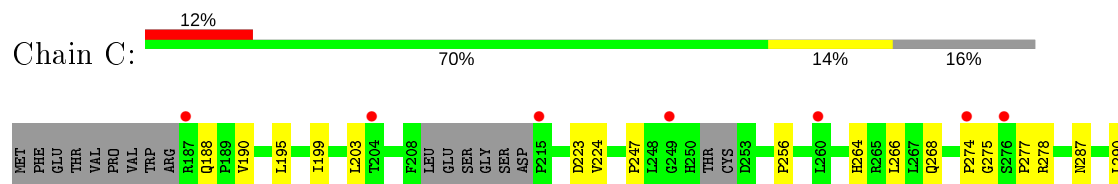
- Molecule 1: DNA (cytosine-5)-methyltransferase 3A

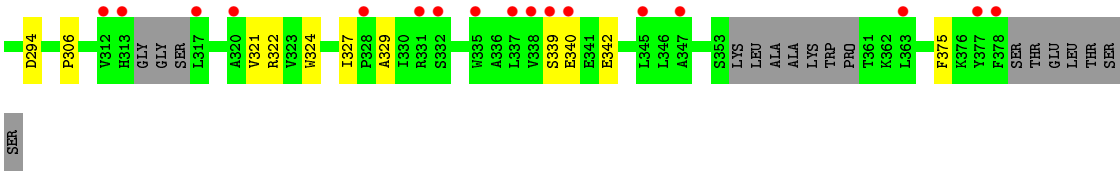


- Molecule 2: DNA (cytosine-5)-methyltransferase 3-like

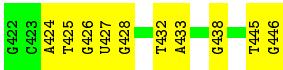


- Molecule 2: DNA (cytosine-5)-methyltransferase 3-like

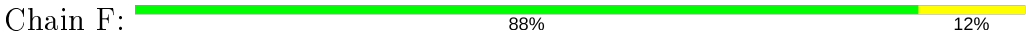




- Molecule 3: DNA (25-MER)



- Molecule 3: DNA (25-MER)



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	205.42Å 205.42Å 89.34Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	33.11 – 2.65 33.11 – 2.65	Depositor EDS
% Data completeness (in resolution range)	98.4 (33.11-2.65) 98.4 (33.11-2.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.93 (at 2.65Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.191 , 0.222 0.193 , 0.223	Depositor DCC
$R_{free}$ test set	2005 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	65.0	Xtriage
Anisotropy	0.205	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 58.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.010 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8380	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, 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	0.24	0/2321	0.39	0/3138
1	D	0.24	0/2323	0.39	0/3141
2	B	0.23	0/1441	0.41	0/1977
2	C	0.25	0/1299	0.43	0/1780
3	E	0.53	0/550	0.90	0/846
3	F	0.53	0/550	0.91	0/846
All	All	0.29	0/8484	0.51	0/11728

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

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	2267	0	2212	30	0
1	D	2269	0	2217	20	0
2	B	1398	0	1235	14	0
2	C	1263	0	1086	16	0
3	E	510	0	279	9	0
3	F	510	0	279	2	0
4	A	24	0	32	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	6	0	8	1	0
4	E	6	0	8	1	0
5	A	26	0	19	1	0
5	D	26	0	19	0	0
6	A	22	0	0	1	0
6	B	5	0	0	0	0
6	C	2	0	0	0	0
6	D	38	0	0	0	0
6	E	1	0	0	0	0
6	F	7	0	0	0	0
All	All	8380	0	7394	81	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 5.

All (81) 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:733:GLU:OE1	1:A:736:ARG:NH1	2.23	0.71
1:A:866:ARG:HE	4:A:1003:GOL:H2	1.56	0.70
1:D:790:ARG:NH1	1:D:892:SER:O	2.28	0.66
1:D:820:GLU:OE1	1:D:866:ARG:NH2	2.29	0.66
1:A:840:ILE:HD11	1:A:859:LEU:HD12	1.79	0.64
2:C:322:ARG:NH2	2:C:340:GLU:OE1	2.22	0.63
2:C:247:PRO:HA	2:C:287:ASN:HD22	1.63	0.63
1:D:733:GLU:OE1	1:D:736:ARG:NH1	2.31	0.63
1:D:690:VAL:O	1:D:736:ARG:NH2	2.31	0.62
1:A:690:VAL:O	1:A:736:ARG:NH2	2.31	0.62
1:D:882:ARG:NH1	3:F:438:DG:OP2	2.33	0.62
1:D:835:THR:HG21	4:E:501:GOL:H31	1.81	0.61
1:A:882:ARG:NH1	3:E:438:DG:OP2	2.34	0.60
1:D:840:ILE:HD11	1:D:859:LEU:HD12	1.83	0.60
1:A:709:PRO:HB3	4:A:1002:GOL:H31	1.82	0.60
1:A:885:ARG:NH2	1:D:876:ASP:OD1	2.34	0.59
1:A:876:ASP:OD1	1:D:885:ARG:NH2	2.35	0.59
1:D:882:ARG:NH2	1:D:886:GLN:OE1	2.35	0.58
2:C:342:GLU:OE1	2:C:342:GLU:N	2.38	0.56
2:C:256:PRO:HB2	2:C:290:LEU:HD23	1.90	0.54
1:A:790:ARG:NH1	1:A:892:SER:O	2.41	0.54
1:D:792:ARG:NH1	3:E:427:PYO:O4'	2.41	0.54
3:E:432:DT:H2''	3:E:433:DA:C8	2.44	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:339:SER:OG	2:C:340:GLU:N	2.42	0.52
2:C:306:PRO:HB3	2:C:324:TRP:CE2	2.45	0.52
2:B:334:HIS:O	2:B:334:HIS:ND1	2.43	0.51
1:A:878:SER:HA	1:D:860:TRP:CE2	2.45	0.51
3:E:424:DA:H2"	3:E:425:DT:H72	1.93	0.49
1:D:835:THR:HG22	1:D:887:ARG:HA	1.94	0.49
2:C:275:GLY:O	2:C:277:PRO:HD3	2.12	0.49
2:C:327:ILE:HG22	2:C:329:ALA:H	1.78	0.48
1:D:875:THR:O	1:D:885:ARG:HD3	2.14	0.48
1:D:710:CYS:N	3:E:427:PYO:HC5	2.28	0.48
2:B:191:ARG:HD2	2:B:237:PRO:O	2.14	0.48
2:C:224:VAL:HB	2:C:266:LEU:HD11	1.96	0.48
4:A:1002:GOL:H11	5:A:1005:SAH:H8	1.97	0.47
1:A:715:ILE:HG12	1:A:760:ALA:HB1	1.98	0.46
2:C:195:LEU:HD22	2:C:224:VAL:HG11	1.97	0.46
1:A:875:THR:O	1:A:885:ARG:HD3	2.16	0.46
3:E:445:DT:H2"	3:E:446:DG:C8	2.51	0.46
1:D:745:GLU:HB2	2:C:274:PRO:HG3	1.97	0.46
1:A:639:LEU:HD12	1:A:734:PHE:CD2	2.51	0.46
2:B:182:VAL:O	2:B:187:ARG:NH1	2.46	0.45
2:C:190:VAL:HB	2:C:375:PHE:CG	2.51	0.45
1:A:808:THR:HG23	1:A:810:ASN:H	1.82	0.45
2:B:229:ARG:O	2:B:233:GLU:HG3	2.17	0.45
2:B:264:HIS:O	2:B:268:GLN:HG2	2.17	0.44
1:D:831:ARG:NH2	3:E:425:DT:H5"	2.33	0.44
2:B:182:VAL:HG22	2:B:183:PRO:HD2	1.98	0.44
1:A:906:LYS:HB2	1:A:911:CYS:SG	2.57	0.44
2:B:308:THR:O	2:B:310:PRO:HD3	2.16	0.44
1:A:743:PRO:HB2	1:A:747:ASP:HB3	1.99	0.44
2:C:264:HIS:O	2:C:268:GLN:HG2	2.17	0.44
3:E:425:DT:H2"	3:E:426:DG:C8	2.53	0.44
1:A:826:LYS:HD3	1:A:851:PHE:CZ	2.53	0.43
2:B:268:GLN:HE22	2:B:271:ARG:HH21	1.66	0.43
2:B:353:SER:O	2:B:353:SER:OG	2.36	0.43
1:D:798:LEU:O	1:D:801:MET:HG3	2.17	0.43
2:C:199:ILE:HB	2:C:203:LEU:HD13	2.00	0.43
1:A:777:PRO:HB3	1:A:795:TRP:CE2	2.53	0.43
1:A:826:LYS:HD3	1:A:851:PHE:CE2	2.54	0.43
2:B:240:LEU:HD23	2:B:281:PHE:HB2	2.00	0.43
1:A:631:ARG:NH1	1:A:906:LYS:O	2.52	0.42
1:A:865:GLU:OE1	1:A:875:THR:OG1	2.33	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:674:MET:HB3	6:A:1117:HOH:O	2.19	0.42
1:A:752:PHE:HB3	1:A:798:LEU:HD23	2.01	0.42
2:B:202:GLU:N	2:B:202:GLU:OE1	2.48	0.42
3:E:428:DG:OP2	3:E:428:DG:H8	2.02	0.42
1:A:813:LEU:O	1:A:829:LYS:NZ	2.52	0.42
1:A:764:SER:OG	1:A:767:ARG:NH2	2.53	0.42
1:A:772:PHE:CE2	2:B:265:ARG:HB2	2.55	0.42
3:F:422:DG:H1'	3:F:423:DC:H5'	2.01	0.42
2:C:190:VAL:HB	2:C:375:PHE:CD2	2.55	0.41
2:B:194:SER:OG	2:B:197:GLU:O	2.37	0.41
1:A:835:THR:HG22	1:A:887:ARG:HA	2.02	0.41
1:D:752:PHE:HB3	1:D:798:LEU:HD23	2.03	0.41
1:A:724:TYR:CZ	2:B:258:TRP:HD1	2.38	0.41
1:D:751:PHE:H	4:D:1001:GOL:H12	1.84	0.41
1:A:808:THR:HG22	1:A:811:ASP:OD2	2.21	0.41
1:A:649:VAL:HG21	1:A:898:ILE:HB	2.03	0.41
2:C:290:LEU:HD22	2:C:294:ASP:HB3	2.02	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	283/285 (99%)	276 (98%)	7 (2%)	0	100	100
1	D	283/285 (99%)	274 (97%)	9 (3%)	0	100	100
2	B	180/208 (86%)	171 (95%)	9 (5%)	0	100	100
2	C	164/208 (79%)	148 (90%)	15 (9%)	1 (1%)	25	37
All	All	910/986 (92%)	869 (96%)	40 (4%)	1 (0%)	51	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	278	ARG

### 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	235/250 (94%)	234 (100%)	1 (0%)	91	95
1	D	238/250 (95%)	238 (100%)	0	100	100
2	B	129/190 (68%)	127 (98%)	2 (2%)	62	78
2	C	114/190 (60%)	111 (97%)	3 (3%)	46	64
All	All	716/880 (81%)	710 (99%)	6 (1%)	81	89

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	877	VAL
2	B	182	VAL
2	B	196	PHE
2	C	188	GLN
2	C	223	ASP
2	C	321	VAL

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

Mol	Chain	Res	Type
2	C	287	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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
3	PYO	E	427	1,3	14,20,21	1.10	1 (7%)	16,28,31	3.32	2 (12%)
3	PYO	F	427	1,3	14,20,21	1.05	1 (7%)	16,28,31	3.58	2 (12%)

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	1,3	-	2/5/25/26	0/2/2/2
3	PYO	F	427	1,3	-	2/5/25/26	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	427	PYO	C2-N3	-3.19	1.31	1.38
3	F	427	PYO	C2-N3	-3.04	1.32	1.38

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	427	PYO	C4-N3-C2	13.83	121.24	114.42
3	E	427	PYO	C4-N3-C2	12.68	120.67	114.42
3	F	427	PYO	C5-C4-N3	-2.47	119.39	123.43
3	E	427	PYO	C5-C4-N3	-2.29	119.68	123.43

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	427	PYO	O4'-C1'-N1-C6
3	E	427	PYO	C2'-C1'-N1-C6
3	F	427	PYO	O4'-C1'-N1-C6
3	F	427	PYO	C2'-C1'-N1-C6

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	427	PYO	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

8 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	GOL	A	1001	-	5,5,5	0.35	0	5,5,5	0.38	0
4	GOL	A	1002	-	5,5,5	0.34	0	5,5,5	0.29	0
4	GOL	D	1001	-	5,5,5	0.34	0	5,5,5	0.34	0
4	GOL	E	501	-	5,5,5	0.36	0	5,5,5	0.37	0
4	GOL	A	1003	-	5,5,5	0.36	0	5,5,5	0.26	0
5	SAH	D	1002	-	21,28,28	1.20	2 (9%)	20,40,40	1.63	2 (10%)
4	GOL	A	1004	-	5,5,5	0.34	0	5,5,5	0.31	0
5	SAH	A	1005	-	21,28,28	1.20	2 (9%)	20,40,40	1.64	2 (10%)

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	GOL	A	1001	-	-	2/4/4/4	-
4	GOL	A	1002	-	-	4/4/4/4	-
4	GOL	D	1001	-	-	2/4/4/4	-
4	GOL	E	501	-	-	0/4/4/4	-
4	GOL	A	1003	-	-	1/4/4/4	-
5	SAH	D	1002	-	-	3/7/31/31	0/3/3/3
4	GOL	A	1004	-	-	3/4/4/4	-
5	SAH	A	1005	-	-	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)
5	A	1005	SAH	C2-N3	4.03	1.38	1.32
5	D	1002	SAH	C2-N3	3.95	1.38	1.32
5	D	1002	SAH	C2-N1	2.45	1.38	1.33
5	A	1005	SAH	C2-N1	2.39	1.38	1.33

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	1002	SAH	N3-C2-N1	-5.64	119.87	128.68
5	A	1005	SAH	N3-C2-N1	-5.53	120.03	128.68
5	A	1005	SAH	C5'-SD-CG	-3.10	92.96	102.27
5	D	1002	SAH	C5'-SD-CG	-2.80	93.88	102.27

There are no chirality outliers.

All (18) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1002	GOL	O1-C1-C2-C3
4	A	1002	GOL	C1-C2-C3-O3
4	D	1001	GOL	C1-C2-C3-O3
5	D	1002	SAH	N-CA-CB-CG
4	A	1004	GOL	O1-C1-C2-C3
5	A	1005	SAH	N-CA-CB-CG
5	A	1005	SAH	C-CA-CB-CG
4	A	1004	GOL	O1-C1-C2-O2
4	A	1001	GOL	C1-C2-C3-O3
4	A	1001	GOL	O2-C2-C3-O3
4	A	1002	GOL	O2-C2-C3-O3
4	D	1001	GOL	O2-C2-C3-O3

*Continued on next page...*



*Continued from previous page...*

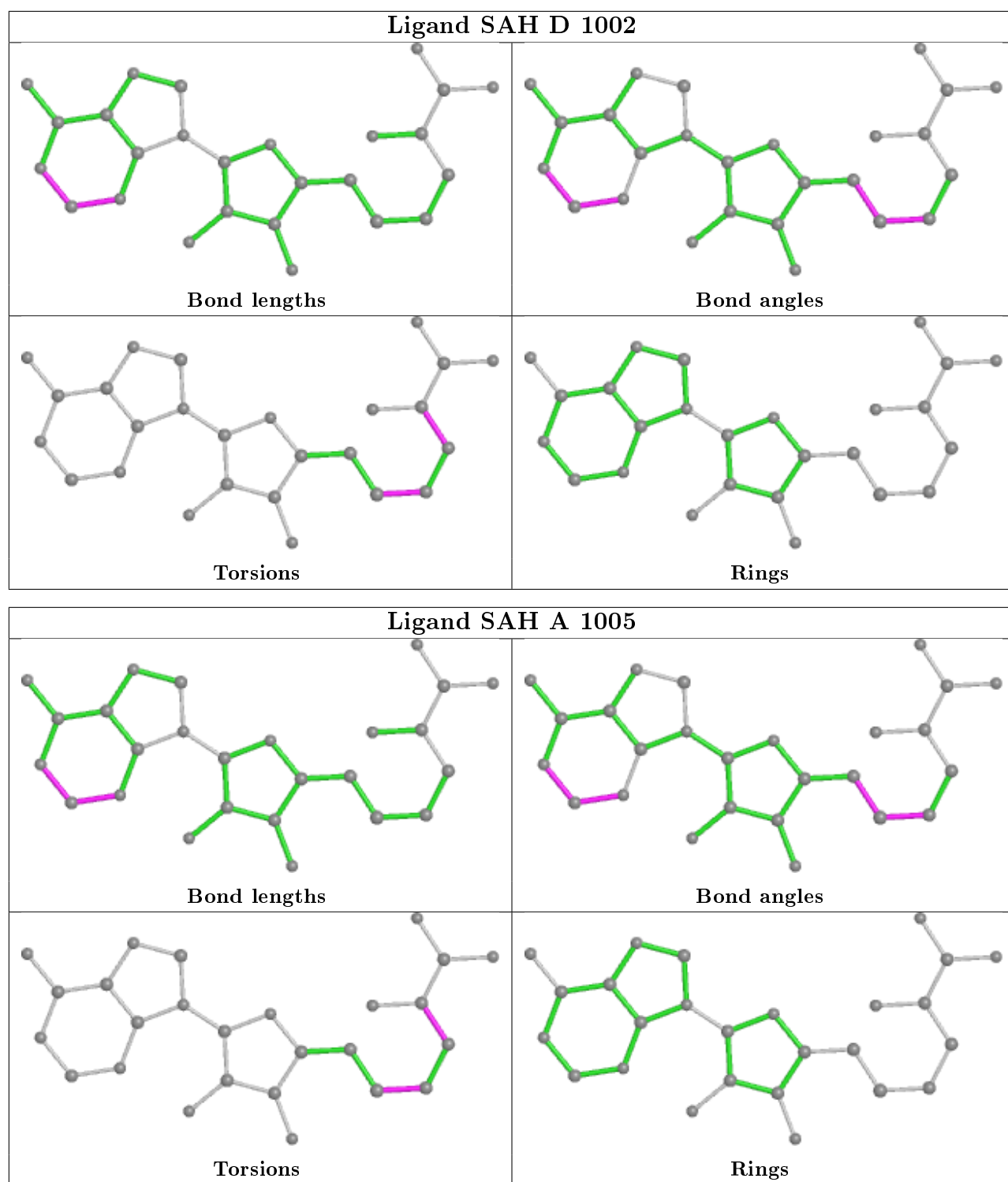
Mol	Chain	Res	Type	Atoms
4	A	1002	GOL	O1-C1-C2-O2
5	D	1002	SAH	CB-CG-SD-C5'
5	A	1005	SAH	CB-CG-SD-C5'
4	A	1003	GOL	C1-C2-C3-O3
4	A	1004	GOL	C1-C2-C3-O3
5	D	1002	SAH	C-CA-CB-CG

There are no ring outliers.

5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1002	GOL	2	0
4	D	1001	GOL	1	0
4	E	501	GOL	1	0
4	A	1003	GOL	1	0
5	A	1005	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 ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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%)	-0.19	5 (1%) 68 65	41, 62, 95, 143	0
1	D	285/285 (100%)	-0.30	3 (1%) 80 79	40, 56, 88, 118	0
2	B	188/208 (90%)	0.38	21 (11%) 5 3	64, 100, 144, 156	0
2	C	174/208 (83%)	0.64	24 (13%) 2 2	72, 122, 161, 172	0
3	E	24/25 (96%)	-0.34	0 100 100	69, 87, 108, 114	0
3	F	24/25 (96%)	-0.37	0 100 100	66, 91, 103, 108	0
All	All	980/1036 (94%)	0.03	53 (5%) 25 23	40, 75, 145, 172	0

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	361	THR	4.9
2	B	317	LEU	4.6
2	B	248	LEU	4.0
2	B	212	GLY	3.9
2	C	274	PRO	3.7
2	C	276	SER	3.5
2	C	317	LEU	3.5
1	A	629	GLU	3.4
2	C	378	PHE	3.4
2	C	335	TRP	3.4
2	B	336	ALA	3.3
2	C	313	HIS	3.1
2	C	331	ARG	3.1
2	C	328	PRO	3.0
2	C	187	ARG	3.0
2	B	213	SER	3.0
2	C	339	SER	2.9
2	C	312	VAL	2.9
2	B	211	SER	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	C	347	ALA	2.9
1	A	804	PRO	2.9
2	C	204	THR	2.9
2	C	332	SER	2.8
2	B	275	GLY	2.6
2	B	344	SER	2.6
2	C	338	VAL	2.5
2	B	222	VAL	2.5
2	C	249	GLY	2.5
2	B	218	LEU	2.5
2	B	192	VAL	2.5
2	B	353	SER	2.5
2	B	338	VAL	2.4
1	D	628	ALA	2.4
2	B	364	VAL	2.4
2	C	345	LEU	2.4
2	B	250	HIS	2.4
2	B	274	PRO	2.3
2	C	320	ALA	2.3
2	B	318	GLN	2.3
2	C	260	LEU	2.3
2	C	377	TYR	2.3
1	D	809	VAL	2.3
2	C	363	LEU	2.2
1	A	805	LEU	2.2
1	A	628	ALA	2.2
2	C	215	PRO	2.2
2	B	319	ASN	2.2
2	B	288	LEU	2.2
2	B	345	LEU	2.1
2	C	340	GLU	2.1
1	D	851	PHE	2.0
1	A	890	GLY	2.0
2	C	337	LEU	2.0

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

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	E	427	19/20	0.96	0.15	59,61,68,68	0
3	PYO	F	427	19/20	0.97	0.17	51,57,62,63	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

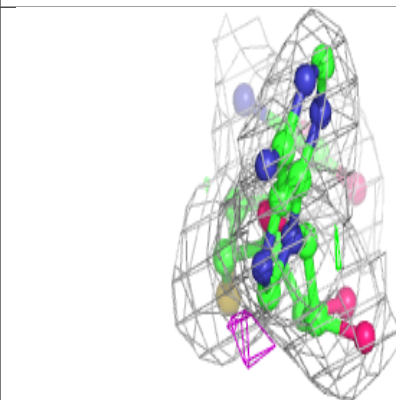
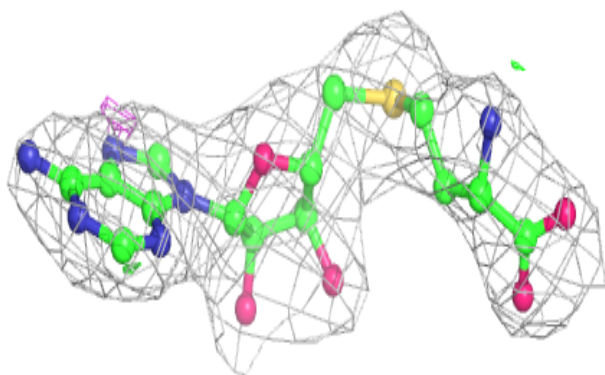
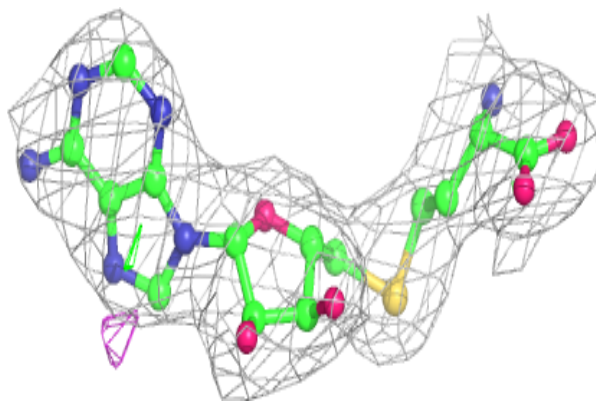
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	GOL	A	1004	6/6	0.73	0.21	69,79,87,90	0
4	GOL	A	1003	6/6	0.75	0.27	76,82,87,89	0
4	GOL	D	1001	6/6	0.86	0.20	83,86,89,90	0
4	GOL	A	1002	6/6	0.87	0.24	65,70,73,82	0
4	GOL	E	501	6/6	0.88	0.43	65,83,88,91	0
4	GOL	A	1001	6/6	0.92	0.51	89,95,100,102	0
5	SAH	A	1005	26/26	0.92	0.22	39,56,68,71	0
5	SAH	D	1002	26/26	0.96	0.18	45,53,63,67	0

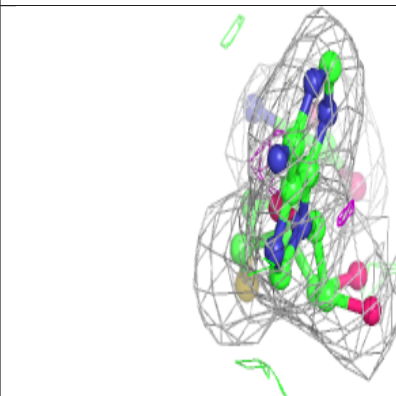
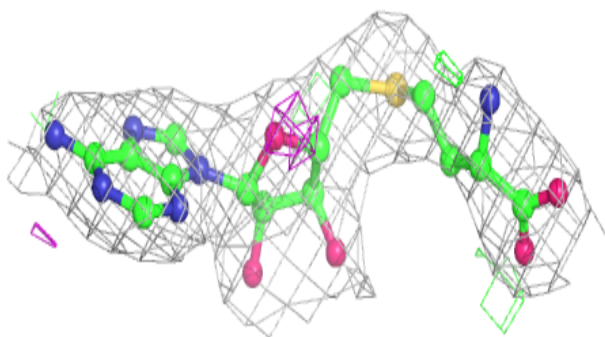
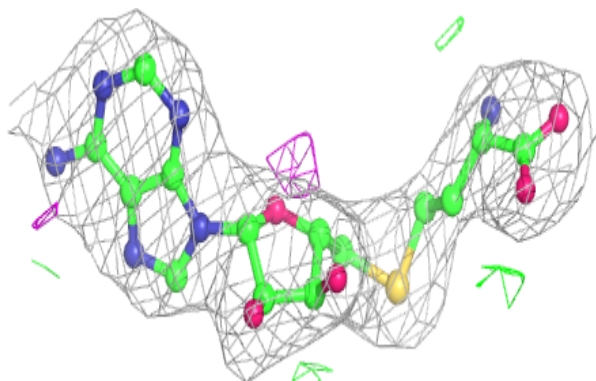
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around SAH A 1005:**

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

**Electron density around SAH D 1002:**

$2mF_o-DF_c$  (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.