



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 3, 2023 – 08:56 AM EDT

PDB ID : 3SV3  
Title : Crystal structure of the large fragment of DNA polymerase I from *Thermus Aquaticus* in a closed ternary complex with the artificial base pair dNaM-d5SICSTP  
Authors : Betz, K.; Diederichs, K.; Marx, A.  
Deposited on : 2011-07-12  
Resolution : 2.10 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 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

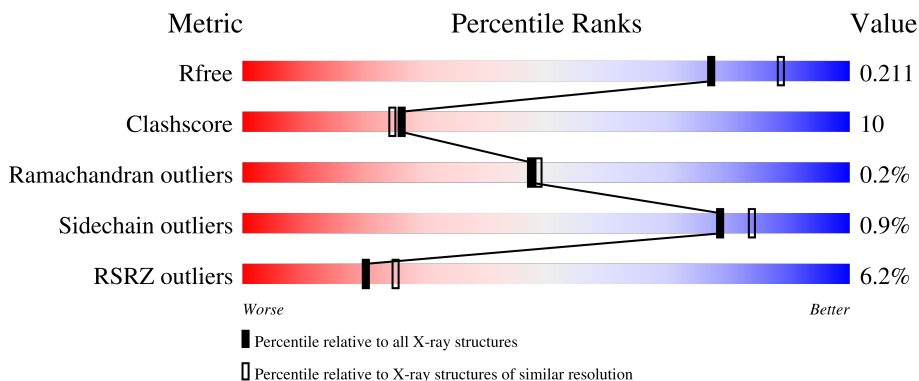
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	540	
2	B	12	
3	C	16	

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
7	FMT	C	217	-	-	X	-

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 9585 atoms, of which 4376 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 polymerase I, thermostable.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	540	8692	2744	4376	773	786	13	0	5	0

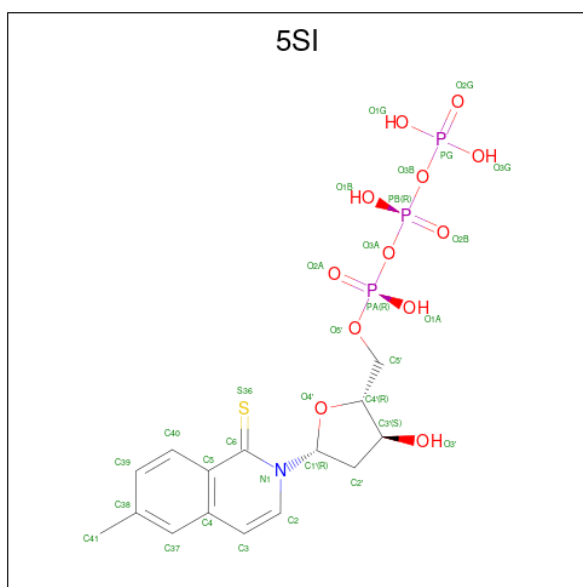
- Molecule 2 is a DNA chain called (5'-D(\*GP\*AP\*CP\*CP\*AP\*CP\*GP\*GP\*CP\*GP\*CP\*(DOC))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	12	240	114	48	67	11	0	0	0

- Molecule 3 is a DNA chain called (5'-D(\*AP\*AP\*AP\*(BMN)P\*GP\*GP\*CP\*GP\*CP\*CP\*GP\*TP\*GP\*GP\*TP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	13	271	132	46	80	13	0	0	0

- Molecule 4 is 2-{2-deoxy-5-O-[(R)-hydroxy{[(R)-hydroxy(phosphonooxy)phosphoryl]oxy}p-hosphoryl]-beta-D-erythro-pentofuranosyl}-6-methylisoquinoline-1(2H)-thione (three-letter code: 5SI) (formula: C<sub>15</sub>H<sub>20</sub>NO<sub>12</sub>P<sub>3</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	N	O	P			S
4	A	1	32	15	1	12	3	1	0	0
4	A	1	32	15	1	12	3	1	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

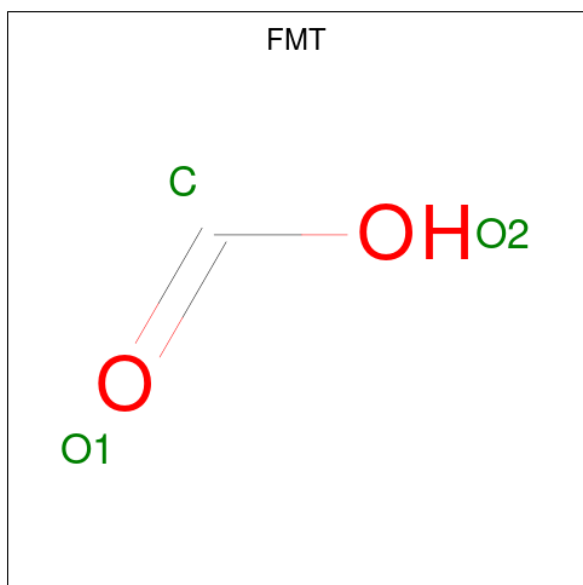
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
5	A	4	4	4	0	0
5	B	1	1	1	0	0

- Molecule 6 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
6	A	1	Total	C	O	0	0
			6	3	3		
6	A	1	Total	C	O	0	0
			6	3	3		
6	C	1	Total	C	O	0	0
			6	3	3		
6	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is FORMIC ACID (three-letter code: FMT) (formula: CH<sub>2</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	A	1	Total C O 3 1 2	0	0
7	C	1	Total C O 3 1 2	0	0
7	C	1	Total C O 3 1 2	0	0
7	C	1	Total C O 3 1 2	0	0
7	C	1	Total C O 3 1 2	0	0
7	C	1	Total C O 3 1 2	0	0

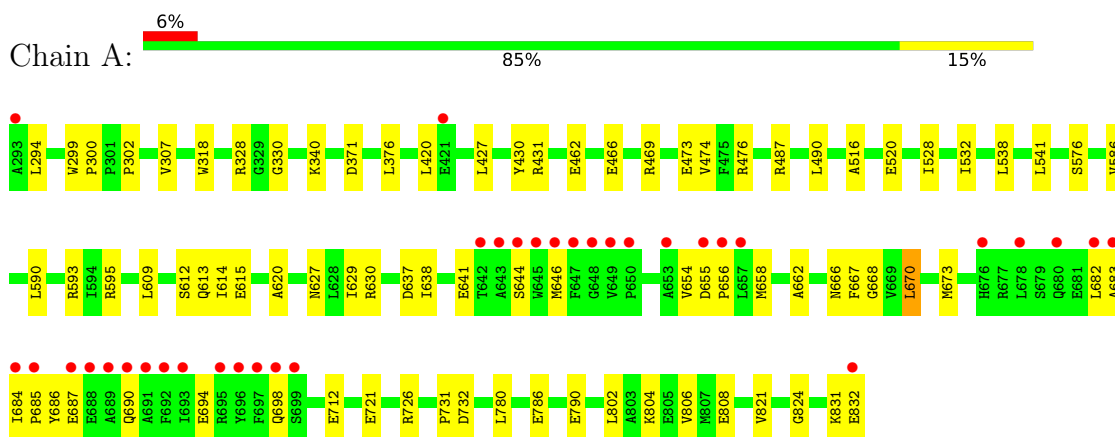
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	197	Total O 197 197	0	0
8	B	22	Total O 22 22	0	0
8	C	22	Total O 22 22	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.

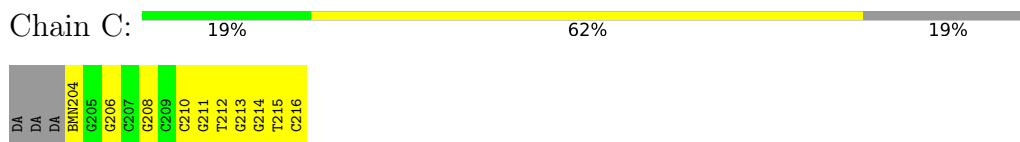
- Molecule 1: DNA polymerase I, thermostable



- Molecule 2: (5'-D(\*GP\*AP\*CP\*CP\*AP\*CP\*GP\*GP\*CP\*GP\*CP\*(DOC))-3')



- Molecule 3: (5'-D(\*AP\*AP\*AP\*(BMN)P\*GP\*GP\*CP\*GP\*CP\*CP\*GP\*TP\*GP\*GP\*TP\*C)-3')





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	108.55Å 108.55Å 90.98Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.00 – 2.10 47.00 – 2.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.00-2.10) 100.0 (47.00-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 2.10Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, $R_{free}$	0.167 , 0.215 0.164 , 0.211	Depositor DCC
$R_{free}$ test set	1816 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.0	Xtrriage
Anisotropy	0.066	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 49.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.037 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9585	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.64% 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: DOC, GOL, 5SI, MG, BMN, FMT

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	0.32	0/4421	0.49	0/5988
2	B	0.76	0/249	1.37	5/382 (1.3%)
3	C	0.73	0/277	1.48	3/426 (0.7%)
All	All	0.39	0/4947	0.67	8/6796 (0.1%)

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	C	212	DT	O4'-C1'-N1	-9.24	101.53	108.00
3	C	206	DG	O4'-C4'-C3'	-7.92	101.25	106.00
2	B	105	DA	O4'-C1'-N9	-6.74	103.28	108.00
2	B	109	DC	O4'-C1'-N1	6.36	112.45	108.00
2	B	108	DG	C1'-O4'-C4'	-6.16	103.94	110.10
2	B	108	DG	O4'-C1'-N9	6.13	112.29	108.00
3	C	216	DC	O4'-C1'-N1	-5.83	103.92	108.00
2	B	106	DC	O4'-C1'-N1	-5.15	104.39	108.00

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	4316	4376	4375	80	0
2	B	240	0	134	5	0
3	C	271	0	151	13	0
4	A	64	0	32	5	0
5	A	4	0	0	0	0
5	B	1	0	0	0	0
6	A	12	0	16	1	0
6	C	12	0	16	3	0
7	A	33	0	11	3	0
7	C	15	0	5	4	0
8	A	197	0	0	8	0
8	B	22	0	0	0	0
8	C	22	0	0	2	0
All	All	5209	4376	4740	101	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 10.

All (101) 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)
3:C:211:DG:H21	7:C:217:FMT:H	1.19	1.05
3:C:211:DG:N7	6:C:218:GOL:H2	1.82	0.93
6:C:218:GOL:H32	8:C:36:HOH:O	1.71	0.90
4:A:1:5SI:O3G	8:A:905:HOH:O	2.00	0.78
1:A:466:GLU:OE1	1:A:469:ARG:NH1	2.19	0.76
3:C:211:DG:N2	7:C:217:FMT:H	2.00	0.75
1:A:613[B]:GLN:NE2	1:A:637:ASP:OD1	2.22	0.72
3:C:214:DG:H2''	3:C:215:DT:H5'	1.75	0.69
1:A:684:ILE:CG2	1:A:685:PRO:HD2	2.24	0.67
3:C:213:DG:H2''	3:C:214:DG:H5''	1.77	0.67
1:A:804:LYS:NZ	4:A:1:5SI:O1G	2.25	0.66
3:C:214:DG:N7	8:C:35:HOH:O	2.26	0.66
7:A:844:FMT:H	8:A:908:HOH:O	1.96	0.65
1:A:694:GLU:HA	1:A:694:GLU:OE1	1.95	0.65
1:A:682:LEU:HB3	1:A:684:ILE:HG12	1.80	0.64
1:A:732:ASP:OD2	8:A:876:HOH:O	2.15	0.64
1:A:427:LEU:HD11	1:A:431:ARG:NH1	2.13	0.64
3:C:211:DG:H21	7:C:217:FMT:C	2.03	0.63
3:C:211:DG:N7	6:C:218:GOL:C2	2.59	0.63
1:A:684:ILE:HG23	1:A:685:PRO:HD2	1.81	0.62
1:A:466:GLU:HG2	1:A:538:LEU:HD21	1.82	0.62

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:427:LEU:HD11	1:A:431:ARG:CZ	2.30	0.61
3:C:210:DC:H2'	3:C:211:DG:C8	2.36	0.61
1:A:541:LEU:HD12	1:A:590:LEU:HD23	1.83	0.60
1:A:615:GLU:OE2	8:A:112:HOH:O	2.17	0.59
1:A:670:LEU:HD22	1:A:670:LEU:O	2.04	0.58
1:A:694:GLU:OE2	1:A:698:GLN:NE2	2.38	0.57
1:A:473:GLU:OE1	1:A:476:ARG:NH2	2.35	0.57
1:A:831:LYS:CA	1:A:832:GLU:HG2	2.35	0.57
1:A:684:ILE:CG2	1:A:685:PRO:CD	2.83	0.56
1:A:376:LEU:HD22	1:A:420:LEU:HD22	1.88	0.56
1:A:655:ASP:HB2	1:A:656:PRO:HD2	1.88	0.55
1:A:627:ASN:OD1	1:A:630:ARG:NH2	2.39	0.55
1:A:302:PRO:HG2	1:A:328:ARG:HD3	1.88	0.55
1:A:307[B]:VAL:HG23	8:A:67:HOH:O	2.07	0.55
3:C:214:DG:H2'	3:C:215:DT:C6	2.41	0.55
2:B:111:DC:H2'	2:B:112:DOC:H6	1.89	0.54
1:A:687:GLU:HA	1:A:690:GLN:HB3	1.88	0.54
1:A:420:LEU:HD23	1:A:427:LEU:HD23	1.90	0.54
1:A:340:LYS:HD2	7:A:8:FMT:C	2.37	0.53
1:A:667:PHE:CE2	4:A:833:5SI:H2'	2.42	0.53
1:A:684:ILE:HG22	1:A:685:PRO:HD2	1.90	0.53
1:A:646:MET:SD	1:A:662:ALA:HA	2.50	0.52
1:A:654:VAL:HA	1:A:658:MET:CE	2.40	0.52
1:A:721:GLU:HA	1:A:726:ARG:O	2.09	0.52
1:A:655:ASP:HB2	1:A:656:PRO:CD	2.41	0.51
1:A:299:TRP:CG	1:A:300:PRO:HA	2.46	0.50
1:A:318:TRP:CZ3	6:A:2:GOL:H12	2.48	0.48
1:A:638:ILE:HG22	1:A:666:ASN:ND2	2.29	0.48
1:A:684:ILE:HG23	1:A:685:PRO:CD	2.44	0.48
1:A:307[A]:VAL:HG13	8:A:67:HOH:O	2.13	0.48
2:B:101:DG:H2'	2:B:102:DA:C8	2.49	0.48
1:A:654:VAL:HA	1:A:658:MET:SD	2.54	0.47
1:A:831:LYS:HA	1:A:832:GLU:HG2	1.96	0.47
1:A:682:LEU:CB	1:A:684:ILE:HG12	2.44	0.47
1:A:684:ILE:HG22	1:A:685:PRO:CD	2.46	0.46
1:A:609:LEU:O	1:A:786:GLU:HA	2.15	0.46
2:B:101:DG:H2''	2:B:102:DA:H5'	1.97	0.46
1:A:682:LEU:N	1:A:683:ALA:HA	2.30	0.45
1:A:516:ALA:O	1:A:520:GLU:HB2	2.16	0.45
1:A:731:PRO:HD2	8:A:176:HOH:O	2.16	0.45
3:C:213:DG:H2''	3:C:214:DG:C5'	2.46	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:654:VAL:HG22	1:A:658:MET:CE	2.47	0.45
1:A:654:VAL:HA	1:A:658:MET:HE1	1.98	0.45
1:A:687:GLU:HA	1:A:687:GLU:OE1	2.17	0.44
4:A:1:5SI:O1B	4:A:1:5SI:O2G	2.34	0.44
1:A:668:GLY:C	1:A:673:MET:HG2	2.38	0.44
1:A:609:LEU:HG	1:A:821:VAL:HG23	1.99	0.44
1:A:614:ILE:HG23	1:A:615:GLU:N	2.33	0.44
1:A:686:TYR:O	1:A:687:GLU:HB2	2.18	0.44
1:A:294:LEU:HD22	1:A:330:GLY:HA2	1.99	0.44
1:A:462[A]:GLU:OE1	1:A:593:ARG:NH2	2.50	0.44
2:B:111:DC:H2'	2:B:112:DOC:C6	2.47	0.44
1:A:462[A]:GLU:CD	1:A:593:ARG:HH22	2.21	0.43
1:A:694:GLU:O	1:A:698:GLN:HG3	2.18	0.43
1:A:595:ARG:HD3	8:A:157:HOH:O	2.18	0.43
1:A:654:VAL:HG13	1:A:658:MET:HB2	2.01	0.43
1:A:804:LYS:NZ	1:A:808:GLU:OE2	2.50	0.43
1:A:831:LYS:C	1:A:832:GLU:HG2	2.39	0.43
1:A:802:LEU:O	1:A:806:VAL:HG23	2.19	0.42
1:A:694:GLU:OE1	1:A:694:GLU:CA	2.63	0.42
1:A:474:VAL:HG13	1:A:528:ILE:HD11	2.02	0.42
1:A:637:ASP:O	1:A:641:GLU:HG3	2.19	0.42
1:A:466:GLU:CG	1:A:538:LEU:HD21	2.48	0.42
1:A:576:SER:O	3:C:208:DG:H4'	2.19	0.42
1:A:620:ALA:HB1	1:A:629:ILE:HG13	2.01	0.42
1:A:824:GLY:HA2	7:A:838:FMT:C	2.50	0.42
1:A:682:LEU:HD12	1:A:684:ILE:HB	2.02	0.42
1:A:694:GLU:HG3	1:A:698:GLN:NE2	2.35	0.41
1:A:780:LEU:HD11	1:A:790:GLU:HB2	2.02	0.41
1:A:371:ASP:HA	1:A:430:TYR:OH	2.20	0.41
2:B:104:DC:H2'	2:B:105:DA:C8	2.55	0.41
3:C:215:DT:O2	7:C:219:FMT:O2	2.38	0.41
1:A:427:LEU:CD1	1:A:431:ARG:NH1	2.82	0.41
1:A:490:LEU:HD11	1:A:532:ILE:HD13	2.03	0.41
1:A:466:GLU:HG3	1:A:538:LEU:CD2	2.51	0.40
1:A:612:SER:OG	1:A:613[B]:GLN:OE1	2.26	0.40
1:A:654:VAL:HG22	1:A:658:MET:HE1	2.02	0.40
1:A:682:LEU:HB2	1:A:684:ILE:N	2.36	0.40
1:A:613[B]:GLN:HA	4:A:833:5SI:O2B	2.21	0.40
1:A:466:GLU:CG	1:A:538:LEU:CD2	2.99	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	543/540 (101%)	527 (97%)	15 (3%)	1 (0%)	47	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	586	VAL

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	446/441 (101%)	442 (99%)	4 (1%)	78	84

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

Mol	Chain	Res	Type
1	A	487	ARG
1	A	644	SER
1	A	670	LEU
1	A	712	GLU

Sometimes 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](#)

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	BMN	C	204	-	22,25,26	1.78	6 (27%)	31,35,38	1.74	6 (19%)
2	DOC	B	112	2,3	16,19,20	3.01	7 (43%)	20,26,29	1.11	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
3	BMN	C	204	-	-	0/9/23/24	0/3/3/3
2	DOC	B	112	2,3	-	0/7/18/19	0/2/2/2

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	112	DOC	O2-C2	6.76	1.36	1.23
2	B	112	DOC	C2-N3	5.50	1.47	1.36
2	B	112	DOC	C5-C4	4.64	1.53	1.42
2	B	112	DOC	C4-N4	4.44	1.44	1.33
3	C	204	BMN	C2'-C3'	-3.65	1.43	1.52
2	B	112	DOC	C6-C5	2.92	1.41	1.35
3	C	204	BMN	O37-C6	2.88	1.41	1.37
2	B	112	DOC	O4'-C1'	-2.51	1.36	1.42
3	C	204	BMN	O3'-C3'	-2.45	1.38	1.43
3	C	204	BMN	O5'-C5'	-2.31	1.39	1.44
3	C	204	BMN	O4'-C4'	-2.17	1.40	1.45

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	112	DOC	C2-N1	2.17	1.44	1.40
3	C	204	BMN	C2-C1	2.08	1.40	1.36

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	204	BMN	O37-C6-C1	4.96	120.57	115.83
3	C	204	BMN	O37-C6-C5	-4.16	119.86	125.24
3	C	204	BMN	C3'-C2'-C1'	3.53	106.09	102.74
2	B	112	DOC	O4'-C1'-C2'	2.99	109.91	106.67
3	C	204	BMN	O4'-C1'-C1	2.95	112.71	109.74
3	C	204	BMN	C2'-C1'-C1	-2.87	110.61	114.72
3	C	204	BMN	C2-C1-C6	2.36	120.41	117.06
2	B	112	DOC	O4'-C1'-N1	2.09	111.60	107.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	112	DOC	2	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 27 ligands modelled in this entry, 5 are monoatomic - leaving 22 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
7	FMT	C	4	-	2,2,2	0.72	0	1,1,1	0.26	0
7	FMT	C	6	-	2,2,2	0.68	0	1,1,1	0.23	0
4	5SI	A	1	5	28,34,34	1.38	4 (14%)	40,53,53	1.31	4 (10%)
6	GOL	C	3	-	5,5,5	0.38	0	5,5,5	0.20	0
7	FMT	A	839	-	2,2,2	0.70	0	1,1,1	0.25	0
7	FMT	A	847	-	2,2,2	0.68	0	1,1,1	0.16	0
7	FMT	A	8	-	2,2,2	0.72	0	1,1,1	0.31	0
7	FMT	A	838	-	2,2,2	0.66	0	1,1,1	0.06	0
7	FMT	A	845	-	2,2,2	0.71	0	1,1,1	0.19	0
7	FMT	A	843	-	2,2,2	0.74	0	1,1,1	0.26	0
4	5SI	A	833	5	28,34,34	1.35	4 (14%)	40,53,53	1.18	5 (12%)
6	GOL	C	218	-	5,5,5	0.34	0	5,5,5	0.14	0
7	FMT	A	846	-	2,2,2	0.65	0	1,1,1	0.15	0
7	FMT	C	1	-	2,2,2	0.70	0	1,1,1	0.20	0
7	FMT	A	9	-	2,2,2	0.71	0	1,1,1	0.23	0
7	FMT	A	841	-	2,2,2	0.66	0	1,1,1	0.07	0
7	FMT	A	844	-	2,2,2	0.70	0	1,1,1	0.19	0
7	FMT	C	217	-	2,2,2	0.69	0	1,1,1	0.19	0
7	FMT	A	840	-	2,2,2	0.73	0	1,1,1	0.30	0
6	GOL	A	2	-	5,5,5	0.36	0	5,5,5	0.26	0
7	FMT	C	219	-	2,2,2	0.72	0	1,1,1	0.25	0
6	GOL	A	835	-	5,5,5	0.34	0	5,5,5	0.48	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
6	GOL	A	835	-	-	0/4/4/4	-
4	5SI	A	833	5	-	4/22/34/34	0/3/3/3
6	GOL	C	218	-	-	2/4/4/4	-
6	GOL	A	2	-	-	2/4/4/4	-
4	5SI	A	1	5	-	6/22/34/34	0/3/3/3
6	GOL	C	3	-	-	4/4/4/4	-

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1	5SI	C5-C4	4.73	1.48	1.41
4	A	833	5SI	C5-C4	4.45	1.47	1.41

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1	5SI	C2-C3	2.95	1.40	1.33
4	A	1	5SI	C6-S36	-2.92	1.60	1.66
4	A	833	5SI	C6-S36	-2.82	1.60	1.66
4	A	833	5SI	C2-C3	2.75	1.39	1.33
4	A	833	5SI	C4-C3	-2.38	1.39	1.44
4	A	1	5SI	C4-C3	-2.27	1.39	1.44

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1	5SI	PB-O3B-PG	-3.46	120.94	132.83
4	A	1	5SI	C4-C5-C6	-3.41	117.33	120.10
4	A	1	5SI	PB-O3A-PA	-3.15	122.01	132.83
4	A	833	5SI	C4-C5-C6	-2.52	118.05	120.10
4	A	833	5SI	C4-C37-C38	-2.35	119.40	122.66
4	A	833	5SI	C2'-C1'-N1	-2.23	108.62	113.77
4	A	833	5SI	O3G-PG-O1G	2.19	116.01	107.64
4	A	1	5SI	C39-C38-C37	2.02	120.99	117.95
4	A	833	5SI	O4'-C1'-N1	2.00	111.44	107.86

There are no chirality outliers.

All (18) torsion outliers are listed below:

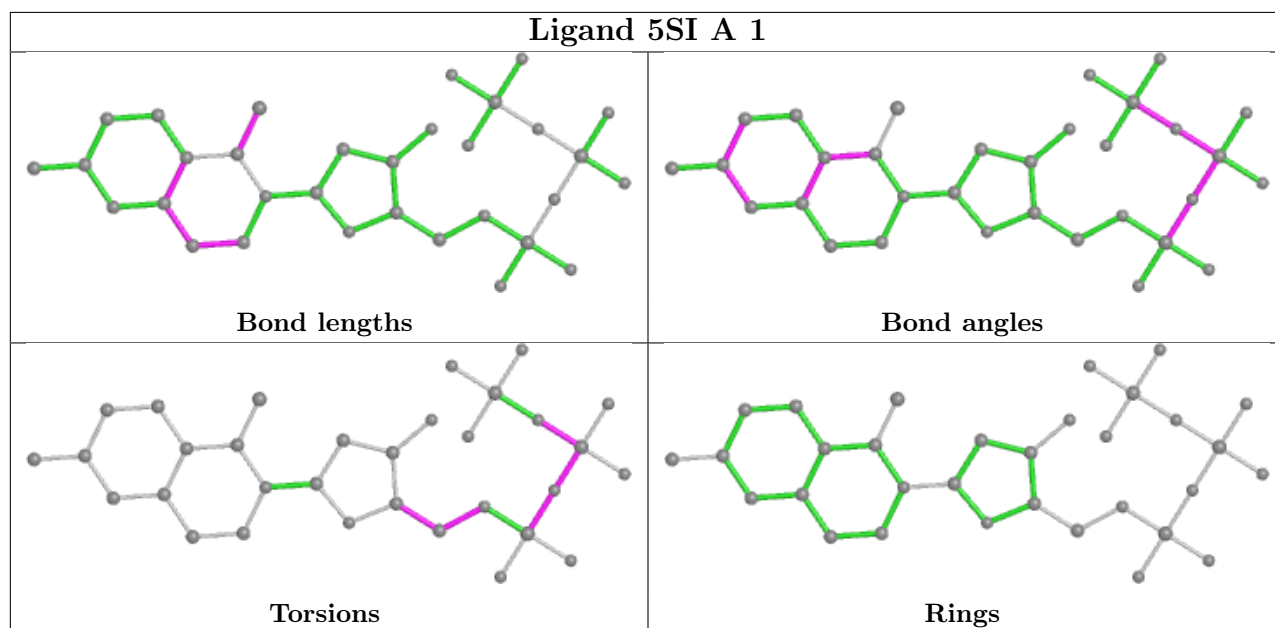
Mol	Chain	Res	Type	Atoms
4	A	1	5SI	C3'-C4'-C5'-O5'
6	A	2	GOL	O1-C1-C2-O2
6	C	3	GOL	O1-C1-C2-C3
6	C	218	GOL	O1-C1-C2-O2
6	C	218	GOL	O1-C1-C2-C3
4	A	1	5SI	O4'-C4'-C5'-O5'
6	A	2	GOL	O1-C1-C2-C3
6	C	3	GOL	C1-C2-C3-O3
6	C	3	GOL	O1-C1-C2-O2
4	A	833	5SI	PG-O3B-PB-O2B
4	A	1	5SI	C4'-C5'-O5'-PA
4	A	1	5SI	PG-O3B-PB-O1B
4	A	833	5SI	PB-O3A-PA-O2A
4	A	833	5SI	PB-O3A-PA-O1A
4	A	833	5SI	PG-O3B-PB-O1B
4	A	1	5SI	PB-O3A-PA-O1A
4	A	1	5SI	PA-O3A-PB-O1B
6	C	3	GOL	O2-C2-C3-O3

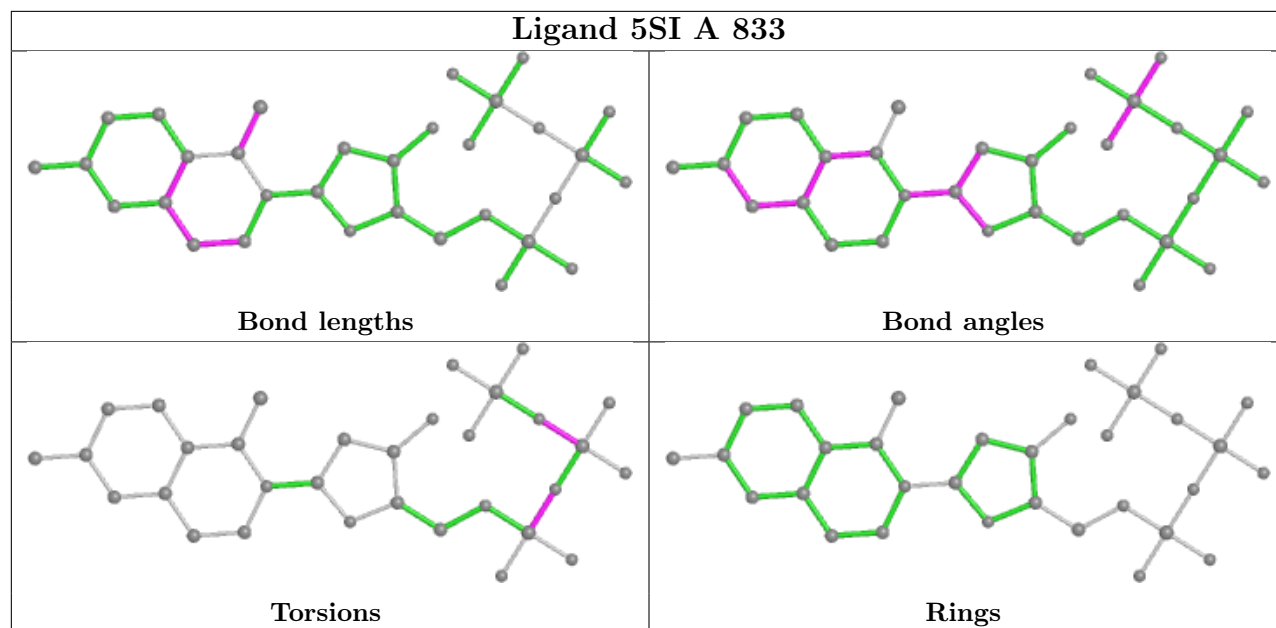
There are no ring outliers.

9 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1	5SI	3	0
7	A	8	FMT	1	0
7	A	838	FMT	1	0
4	A	833	5SI	2	0
6	C	218	GOL	3	0
7	A	844	FMT	1	0
7	C	217	FMT	3	0
6	A	2	GOL	1	0
7	C	219	FMT	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

### 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	540/540 (100%)	0.22	35 (6%) 18 23	26, 42, 92, 130	0
2	B	11/12 (91%)	-0.70	0 100 100	30, 32, 61, 63	0
3	C	12/16 (75%)	-0.54	0 100 100	28, 32, 54, 57	0
All	All	563/568 (99%)	0.18	35 (6%) 20 25	26, 42, 91, 130	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	682	LEU	9.6
1	A	648	GLY	9.4
1	A	692	PHE	9.3
1	A	645	TRP	8.8
1	A	693	ILE	8.7
1	A	649	VAL	8.6
1	A	689	ALA	8.0
1	A	647	PHE	6.4
1	A	650	PRO	5.8
1	A	644	SER	5.3
1	A	696	TYR	5.2
1	A	697	PHE	5.1
1	A	691	ALA	4.7
1	A	688	GLU	4.6
1	A	293	ALA	4.5
1	A	646	MET	4.4
1	A	683	ALA	4.3
1	A	657	LEU	3.9
1	A	653	ALA	3.8
1	A	680	GLN	3.8
1	A	678	LEU	3.6
1	A	695	ARG	3.6
1	A	421	GLU	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	699	SER	2.7
1	A	676	HIS	2.6
1	A	684	ILE	2.6
1	A	656	PRO	2.6
1	A	642	THR	2.6
1	A	643	ALA	2.5
1	A	687	GLU	2.5
1	A	698	GLN	2.4
1	A	690	GLN	2.2
1	A	832	GLU	2.1
1	A	685	PRO	2.1
1	A	655	ASP	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(Å <sup>2</sup> )	Q<0.9
3	BMN	C	204	23/24	0.93	0.11	40,65,98,100	0
2	DOC	B	112	18/19	0.98	0.12	27,35,36,38	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(Å <sup>2</sup> )	Q<0.9
7	FMT	A	846	3/3	0.56	0.22	62,62,64,69	0
7	FMT	A	838	3/3	0.74	0.19	47,47,61,64	0
7	FMT	A	841	3/3	0.75	0.15	41,41,58,58	0
6	GOL	C	218	6/6	0.75	0.27	50,52,57,64	6

*Continued on next page...*

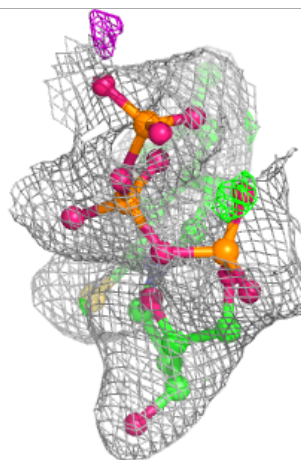
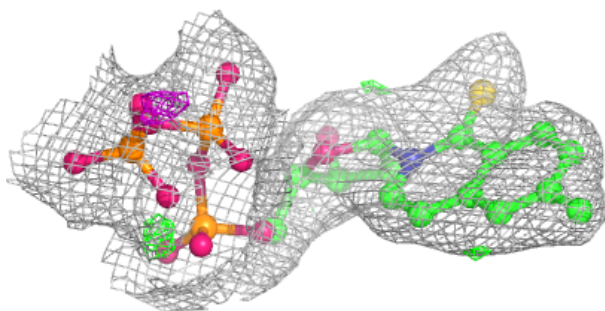
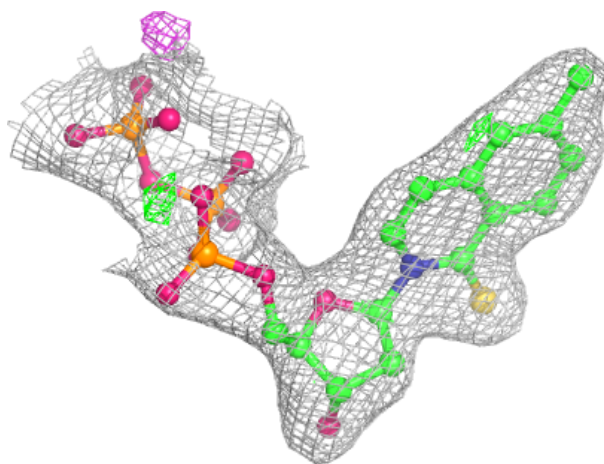
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	FMT	A	844	3/3	0.77	0.34	70,70,74,74	0
7	FMT	A	845	3/3	0.78	0.17	71,71,76,77	0
7	FMT	A	847	3/3	0.81	0.19	61,61,66,67	0
7	FMT	C	6	3/3	0.82	0.16	56,56,61,62	0
5	MG	A	837	1/1	0.83	0.13	51,51,51,51	0
5	MG	A	836	1/1	0.85	0.10	49,49,49,49	0
5	MG	B	1	1/1	0.87	0.05	52,52,52,52	0
7	FMT	C	1	3/3	0.87	0.07	61,61,63,67	0
7	FMT	A	9	3/3	0.88	0.16	60,60,64,65	0
7	FMT	A	839	3/3	0.89	0.27	60,60,65,66	0
7	FMT	A	843	3/3	0.90	0.09	77,77,78,79	0
5	MG	A	842	1/1	0.91	0.05	61,61,61,61	0
7	FMT	A	840	3/3	0.92	0.36	49,49,61,68	0
7	FMT	A	8	3/3	0.93	0.21	62,62,63,64	0
6	GOL	C	3	6/6	0.94	0.13	49,56,60,69	0
4	5SI	A	1	32/32	0.94	0.11	29,60,88,91	0
6	GOL	A	2	6/6	0.94	0.20	53,63,68,76	0
7	FMT	C	219	3/3	0.94	0.11	66,66,72,76	0
7	FMT	C	4	3/3	0.95	0.15	55,55,57,57	0
7	FMT	C	217	3/3	0.96	0.12	47,47,63,71	0
6	GOL	A	835	6/6	0.96	0.10	29,37,42,42	0
4	5SI	A	833	32/32	0.98	0.10	39,48,56,67	0
5	MG	A	834	1/1	0.98	0.15	61,61,61,61	0

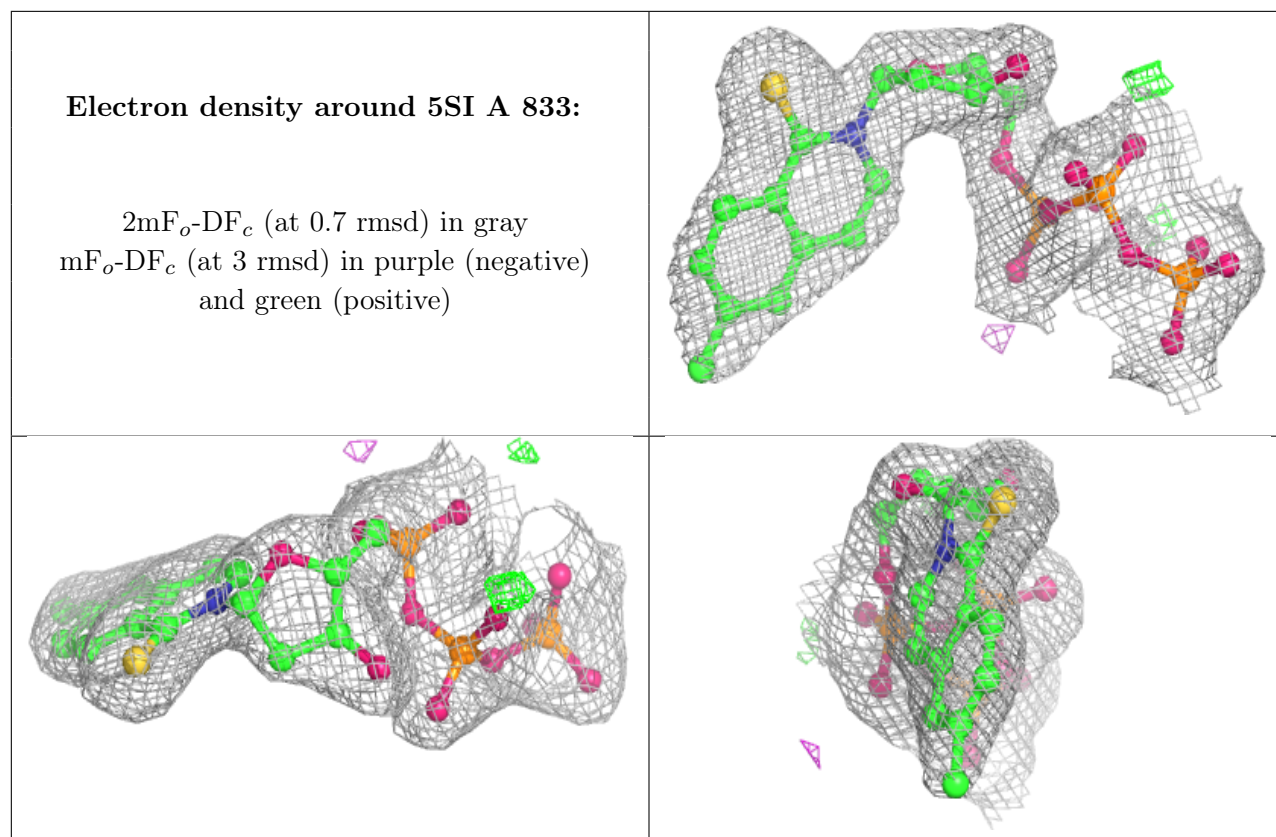
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around 5SI A 1:**

$2mF_o-DF_c$  (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.