



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 3, 2023 – 10:37 AM EDT

PDB ID : 3T3F  
Title : Ternary Structure of the large fragment of Taq DNA polymerase bound to an abasic site and dNTP  
Authors : Marx, A.; Diederichs, K.; Obeid, S.  
Deposited on : 2011-07-25  
Resolution : 1.90 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

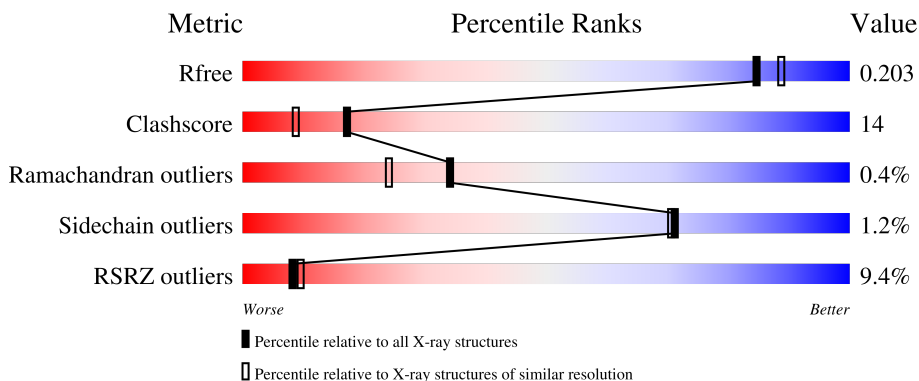
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


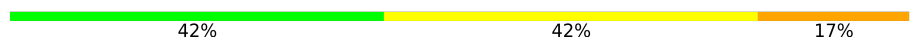
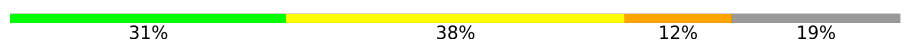
The reported resolution of this entry is 1.90 Å.

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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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
2	DOC	B	112	-	-	X	-
4	GOL	A	13	-	-	X	-
4	GOL	C	2	-	-	X	-
7	N5P	A	836	-	-	X	-
8	FMT	A	2	-	-	X	-

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 5171 atoms, of which 9 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	N	O	S			
1	A	533	4218	2675	761	769	13	0	8	0

- Molecule 2 is a DNA chain called 5'-D(\*GP\*AP\*CP\*CP\*AP\*CP\*GP\*GP\*CP\*GP\*CP\*(D OC))-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\*(3DR)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	259	121	46	79	13	0	0	0

- 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 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0

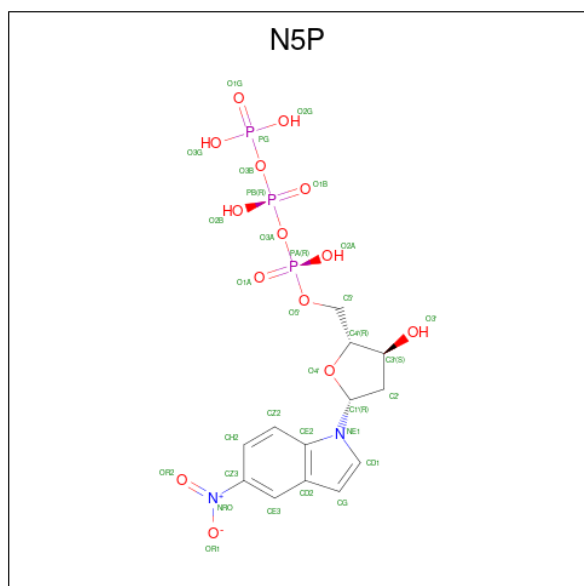
- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	2	Total Mg 2 2	0	0
5	B	1	Total Mg 1 1	0	0
5	C	1	Total Mg 1 1	0	0

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Na 1 1	0	0

- Molecule 7 is 1-{2-DEOXY-5-O-[(R)-HYDROXY{[(R)-HYDROXY(PHOSPHONOOXY)P HOSPHORYL]OXY}PHOSPHORYL]-BETA-D-ERYTHRO-PENTOFURANOSYL}-5-NITRO -1H-INDOLE (three-letter code: N5P) (formula: C<sub>13</sub>H<sub>17</sub>N<sub>2</sub>O<sub>14</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C N O P 32 13 2 14 3	0	0

- Molecule 8 is FORMIC ACID (three-letter code: FMT) (formula: CH<sub>2</sub>O<sub>2</sub>).



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

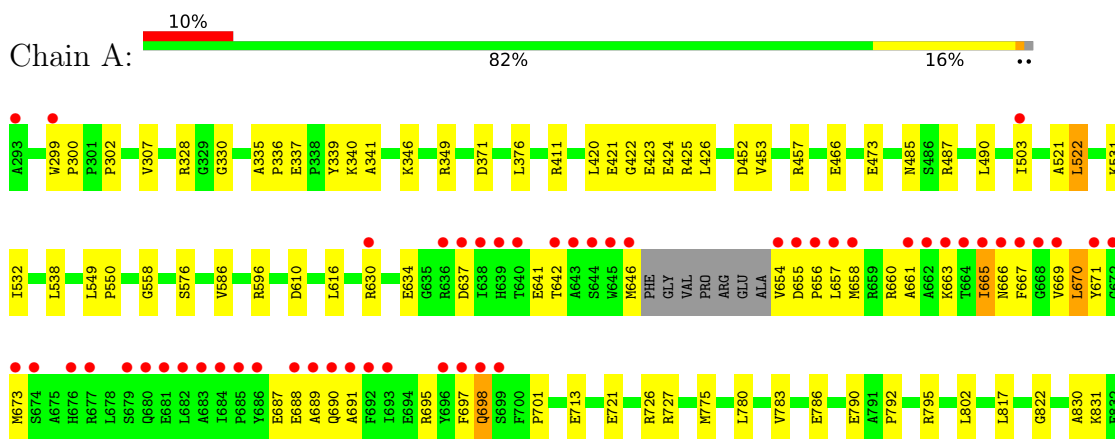
- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	247	Total O 247 247	0	0
9	B	28	Total H O 32 4 28	0	0
9	C	43	Total H O 48 5 43	0	0

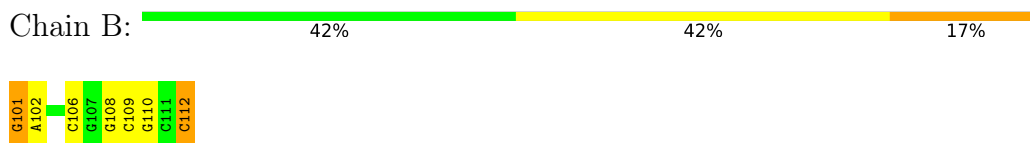
### 3 Residue-property plots

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.

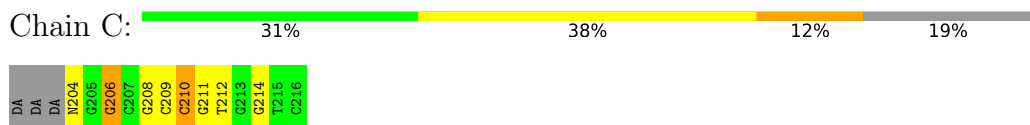
- Molecule 1: DNA polymerase I, thermostable



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



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





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	109.78Å 109.78Å 91.22Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.16 – 1.90 47.54 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (42.16-1.90) 99.8 (47.54-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.56 (at 1.90Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, $R_{free}$	0.170 , 0.206 0.168 , 0.203	Depositor DCC
$R_{free}$ test set	2525 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.5	Xtrriage
Anisotropy	0.224	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 61.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.033 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	5171	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.04% 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: MG, NA, DOC, FMT, N5P, GOL, 3DR

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	0.40	0/4318	0.53	0/5848
2	B	1.04	0/249	1.50	5/382 (1.3%)
3	C	1.02	1/277 (0.4%)	1.55	5/426 (1.2%)
All	All	0.51	1/4844 (0.0%)	0.73	10/6656 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	210	DC	C3'-O3'	-5.03	1.37	1.44

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	212	DT	O4'-C1'-N1	-9.13	101.61	108.00
2	B	108	DG	O4'-C1'-N9	7.56	113.29	108.00
3	C	214	DG	C1'-O4'-C4'	-7.23	102.87	110.10
3	C	206	DG	O4'-C4'-C3'	-6.88	101.75	104.50
3	C	214	DG	O4'-C4'-C3'	-6.21	102.02	104.50
2	B	106	DC	O4'-C1'-N1	-5.96	103.83	108.00
2	B	109	DC	O4'-C1'-N1	5.92	112.14	108.00
2	B	101	DG	O4'-C1'-C2'	5.72	110.48	105.90
3	C	209	DC	O4'-C1'-N1	-5.72	103.99	108.00
2	B	110	DG	O4'-C1'-N9	5.40	111.78	108.00

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	4218	0	4239	99	0
2	B	240	0	134	14	0
3	C	259	0	143	8	0
4	A	42	0	56	11	0
4	B	6	0	8	2	0
4	C	30	0	40	9	0
5	A	2	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
6	A	1	0	0	0	0
7	A	32	0	13	21	0
8	A	12	0	4	3	0
9	A	247	0	0	12	1
9	B	28	4	0	3	0
9	C	43	5	0	2	0
All	All	5162	9	4637	132	1

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 14.

All (132) 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:655:ASP:HB2	1:A:656:PRO:HD2	1.31	1.06
1:A:339:TYR:HB2	8:A:2:FMT:H	1.41	1.02
1:A:689:ALA:HB3	1:A:690:GLN:CB	1.92	1.00
7:A:836:N5P:H5'1	2:B:112:DOC:H2'	1.44	0.99
7:A:836:N5P:H3'	9:A:223:HOH:O	1.64	0.96
1:A:673:MET:HA	3:C:204:3DR:H5''	1.49	0.95
4:A:3:GOL:H2	9:A:62:HOH:O	1.65	0.95
4:C:2:GOL:H11	9:C:217:HOH:O	1.66	0.94
1:A:336:PRO:HG2	4:A:13:GOL:H32	1.53	0.90
1:A:831:LYS:HD3	9:A:219:HOH:O	1.73	0.87
1:A:689:ALA:HB3	1:A:691:ALA:N	1.88	0.87
1:A:689:ALA:CB	1:A:690:GLN:CB	2.54	0.85

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:206:DG:N7	4:C:2:GOL:H12	1.94	0.83
7:A:836:N5P:C5'	7:A:836:N5P:HD1	2.09	0.83
7:A:836:N5P:H5'1	2:B:112:DOC:C2'	2.07	0.83
1:A:670:LEU:C	1:A:670:LEU:HD13	2.00	0.82
1:A:349[A]:ARG:NH1	1:A:371:ASP:OD2	2.13	0.82
1:A:336:PRO:HG2	4:A:13:GOL:C3	2.10	0.81
1:A:421:GLU:HA	9:A:251:HOH:O	1.79	0.80
9:B:116:HOH:O	4:C:5:GOL:H2	1.87	0.74
1:A:339:TYR:HB2	8:A:2:FMT:C	2.17	0.74
1:A:695:ARG:HA	1:A:698:GLN:HB2	1.71	0.73
7:A:836:N5P:C5'	2:B:112:DOC:H2'	2.18	0.72
1:A:654:VAL:HA	1:A:658:MET:SD	2.29	0.72
1:A:610[B]:ASP:OD2	1:A:831:LYS:NZ	2.22	0.72
1:A:655:ASP:HB2	1:A:656:PRO:CD	2.14	0.72
7:A:836:N5P:HD1	7:A:836:N5P:H5'2	1.72	0.72
1:A:689:ALA:HB3	1:A:690:GLN:CA	2.19	0.71
1:A:642:THR:O	1:A:646:MET:HG2	1.93	0.68
1:A:376:LEU:HD22	1:A:420:LEU:HD23	1.75	0.68
7:A:836:N5P:H5'1	7:A:836:N5P:HD1	1.76	0.68
1:A:689:ALA:HB3	1:A:691:ALA:H	1.58	0.67
1:A:302:PRO:HG2	1:A:328:ARG:HD3	1.78	0.65
1:A:307:VAL:HG13	9:A:215:HOH:O	1.95	0.64
1:A:424:GLU:H	1:A:424:GLU:CD	2.01	0.64
1:A:336:PRO:CG	4:A:13:GOL:H32	2.28	0.63
1:A:665:ILE:HG22	1:A:666:ASN:N	2.12	0.63
1:A:689:ALA:HB3	1:A:690:GLN:C	2.19	0.63
7:A:836:N5P:C5'	2:B:112:DOC:C2'	2.76	0.63
1:A:670:LEU:HD13	1:A:670:LEU:O	1.97	0.63
1:A:688:GLU:CB	1:A:691:ALA:HB2	2.29	0.61
1:A:337:GLU:HB3	8:A:2:FMT:O2	2.02	0.60
1:A:670:LEU:CD1	1:A:671:TYR:CE1	2.85	0.59
1:A:630:ARG:O	1:A:634:GLU:HG2	2.02	0.59
7:A:836:N5P:C4'	2:B:112:DOC:C2'	2.80	0.59
1:A:687:GLU:CB	1:A:688:GLU:HA	2.32	0.59
1:A:670:LEU:HD12	1:A:671:TYR:CE2	2.37	0.58
7:A:836:N5P:HH2	9:A:229:HOH:O	2.03	0.58
1:A:822:GLY:HA3	1:A:830:ALA:O	2.03	0.58
1:A:660:ARG:NH1	2:B:112:DOC:OP1	2.35	0.58
1:A:558:GLY:O	9:A:278:HOH:O	2.16	0.57
1:A:670:LEU:CD1	1:A:671:TYR:CD1	2.87	0.57
7:A:836:N5P:C4'	2:B:112:DOC:H2''	2.34	0.57

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:836:N5P:H5'2	7:A:836:N5P:CD1	2.35	0.56
1:A:411:ARG:NH2	4:A:9:GOL:O3	2.37	0.56
1:A:485:ASN:O	4:C:5:GOL:H12	2.06	0.56
1:A:689:ALA:CB	1:A:690:GLN:CA	2.83	0.56
1:A:670:LEU:HD12	1:A:671:TYR:CZ	2.41	0.55
4:C:6:GOL:O2	9:C:50:HOH:O	2.08	0.55
1:A:503:ILE:HD11	1:A:522:LEU:HD13	1.89	0.55
1:A:503:ILE:HD11	1:A:522:LEU:CD1	2.37	0.55
1:A:667:PHE:CE2	7:A:836:N5P:H2'1	2.42	0.55
1:A:490:LEU:HD11	1:A:532:ILE:HD13	1.88	0.54
2:B:101:DG:H2''	2:B:102:DA:H5'	1.89	0.54
1:A:346:LYS:N	1:A:346:LYS:HD2	2.22	0.53
1:A:473:GLU:HG2	1:A:531:LYS:HE2	1.90	0.53
1:A:690:GLN:O	1:A:691:ALA:HB3	2.09	0.53
1:A:663:LYS:O	1:A:667:PHE:HB2	2.09	0.53
1:A:689:ALA:HB1	1:A:690:GLN:CB	2.39	0.52
3:C:206:DG:OP2	4:C:2:GOL:H32	2.10	0.52
1:A:783:VAL:HB	1:A:786[B]:GLU:CG	2.38	0.52
7:A:836:N5P:H5'1	2:B:112:DOC:C3'	2.40	0.52
1:A:655:ASP:OD1	1:A:657:LEU:HB2	2.09	0.52
1:A:670:LEU:O	1:A:670:LEU:HD22	2.10	0.52
7:A:836:N5P:C5'	7:A:836:N5P:CD1	2.85	0.51
1:A:336:PRO:HG2	4:A:13:GOL:O3	2.11	0.51
1:A:420:LEU:HD11	1:A:426:LEU:CB	2.41	0.51
7:A:836:N5P:H4'	2:B:112:DOC:H2''	1.92	0.50
3:C:206:DG:OP2	4:C:2:GOL:C3	2.60	0.50
1:A:299:TRP:CG	1:A:300:PRO:HA	2.47	0.49
1:A:669:VAL:HG23	1:A:673:MET:CE	2.42	0.49
1:A:340:LYS:HB3	4:A:13:GOL:H11	1.95	0.49
1:A:596:ARG:NE	9:A:866:HOH:O	2.45	0.49
1:A:670:LEU:HD12	1:A:671:TYR:CD2	2.48	0.49
1:A:670:LEU:HD11	1:A:671:TYR:CE1	2.48	0.49
1:A:670:LEU:C	1:A:670:LEU:CD1	2.75	0.48
1:A:616:LEU:HG	1:A:667:PHE:HE1	1.79	0.48
7:A:836:N5P:C4'	2:B:112:DOC:H2'	2.43	0.48
1:A:452:ASP:OD2	4:A:11:GOL:O3	2.28	0.47
1:A:775:MET:SD	1:A:802:LEU:HD23	2.54	0.47
1:A:330:GLY:N	9:A:248:HOH:O	2.31	0.47
1:A:423:GLU:OE1	1:A:727:ARG:NH2	2.48	0.47
1:A:425[A]:ARG:NH2	4:A:10:GOL:O3	2.49	0.46
1:A:689:ALA:H	1:A:691:ALA:HB2	1.80	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:697:PHE:O	1:A:701:PRO:HA	2.14	0.46
7:A:836:N5P:C3'	9:A:223:HOH:O	2.41	0.46
1:A:661:ALA:O	1:A:665:ILE:HD13	2.15	0.46
1:A:376:LEU:CD2	1:A:420:LEU:HD23	2.44	0.45
1:A:792:PRO:HG2	1:A:795:ARG:HD2	1.96	0.45
1:A:713:GLU:HB3	9:A:76:HOH:O	2.16	0.45
1:A:420:LEU:HD11	1:A:426:LEU:HB3	1.98	0.45
1:A:466:GLU:HG2	1:A:538:LEU:HD21	1.99	0.45
1:A:663:LYS:NZ	7:A:836:N5P:O2G	2.48	0.45
1:A:335:ALA:HB1	1:A:341:ALA:HB2	1.99	0.44
1:A:642:THR:HG21	1:A:665:ILE:HG21	1.99	0.44
3:C:208:DG:N7	4:C:6:GOL:H31	2.33	0.44
1:A:655:ASP:CB	1:A:656:PRO:HD2	2.20	0.44
2:B:101:DG:H2'	2:B:102:DA:C8	2.52	0.44
1:A:783:VAL:HB	1:A:786[B]:GLU:HG3	2.00	0.43
4:B:12:GOL:H12	9:B:136:HOH:O	2.16	0.43
3:C:208:DG:C8	4:C:6:GOL:H31	2.54	0.43
1:A:721:GLU:HA	1:A:726:ARG:O	2.18	0.43
3:C:210:DC:H2'	3:C:211:DG:C8	2.54	0.43
1:A:688:GLU:H	1:A:691:ALA:HB3	1.84	0.42
1:A:422:GLY:N	9:A:251:HOH:O	2.47	0.42
1:A:596:ARG:HD3	4:A:11:GOL:O1	2.20	0.42
1:A:453:VAL:O	1:A:457:ARG:HG2	2.20	0.42
1:A:503:ILE:CD1	1:A:521:ALA:HB3	2.49	0.42
7:A:836:N5P:O4'	2:B:112:DOC:H2'	2.20	0.42
1:A:780:LEU:HD11	1:A:790:GLU:HB2	2.00	0.42
1:A:549:LEU:N	1:A:550:PRO:CD	2.83	0.42
1:A:420:LEU:HD11	1:A:426:LEU:HB2	2.01	0.42
1:A:422:GLY:O	1:A:424:GLU:OE1	2.37	0.41
1:A:637:ASP:O	1:A:641:GLU:HG3	2.20	0.41
1:A:576:SER:O	3:C:208:DG:H4'	2.20	0.41
1:A:689:ALA:N	1:A:691:ALA:HB2	2.35	0.41
1:A:689:ALA:H	1:A:691:ALA:CB	2.33	0.41
7:A:836:N5P:O4'	2:B:112:DOC:C2'	2.69	0.41
4:B:12:GOL:C1	9:B:136:HOH:O	2.68	0.41
1:A:340:LYS:HG2	4:A:13:GOL:C1	2.51	0.40
1:A:503:ILE:CD1	1:A:522:LEU:HD13	2.51	0.40
1:A:786[A]:GLU:OE2	1:A:831:LYS:HD2	2.22	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
9:A:865:HOH:O	9:A:865:HOH:O[6_555]	2.06	0.14

### 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	537/540 (99%)	520 (97%)	15 (3%)	2 (0%)	34 24

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	586	VAL
1	A	665	ILE

#### 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	427/441 (97%)	422 (99%)	5 (1%)	71 70

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	487	ARG
1	A	522	LEU
1	A	670	LEU
1	A	698	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	817	LEU

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	3DR	C	204	3	8,11,12	0.57	0	9,14,17	0.92	0
2	DOC	B	112	2,3	16,19,20	0.97	1 (6%)	20,26,29	0.86	1 (5%)

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	3DR	C	204	3	-	1/3/15/16	0/1/1/1
2	DOC	B	112	2,3	-	0/7/18/19	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	112	DOC	C6-C5	2.48	1.40	1.35

All (1) bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	112	DOC	C3'-C2'-C1'	2.44	105.60	102.78

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	204	3DR	O4'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	204	3DR	1	0
2	B	112	DOC	12	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 23 ligands modelled in this entry, 5 are monoatomic - leaving 18 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	$\# Z  > 2$
4	GOL	C	6	-	5,5,5	0.33	0	5,5,5	0.38	0
8	FMT	A	838	-	2,2,2	0.71	0	1,1,1	0.24	0
8	FMT	A	837	-	2,2,2	0.67	0	1,1,1	0.19	0
4	GOL	B	12	-	5,5,5	0.40	0	5,5,5	0.36	0
4	GOL	C	8	-	5,5,5	0.42	0	5,5,5	0.51	0
4	GOL	C	5	-	5,5,5	0.48	0	5,5,5	0.13	0
4	GOL	A	10	-	5,5,5	0.32	0	5,5,5	0.46	0
8	FMT	A	2	-	2,2,2	0.76	0	1,1,1	0.29	0
7	N5P	A	836	5	28,34,34	1.22	3 (10%)	38,53,53	1.23	4 (10%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	GOL	C	7	-	5,5,5	0.35	0	5,5,5	0.26	0
4	GOL	A	3	-	5,5,5	0.25	0	5,5,5	0.77	0
4	GOL	A	1	-	5,5,5	0.30	0	5,5,5	0.38	0
4	GOL	A	11	-	5,5,5	0.28	0	5,5,5	0.56	0
4	GOL	A	13	-	5,5,5	0.37	0	5,5,5	0.40	0
4	GOL	C	2	-	5,5,5	0.30	0	5,5,5	0.33	0
8	FMT	A	839	-	2,2,2	0.70	0	1,1,1	0.30	0
4	GOL	A	4	-	5,5,5	0.37	0	5,5,5	0.35	0
4	GOL	A	9	-	5,5,5	0.33	0	5,5,5	0.65	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
4	GOL	C	6	-	-	4/4/4/4	-
4	GOL	B	12	-	-	1/4/4/4	-
4	GOL	C	8	-	-	2/4/4/4	-
4	GOL	C	5	-	-	2/4/4/4	-
4	GOL	A	10	-	-	2/4/4/4	-
7	N5P	A	836	5	-	2/20/38/38	0/3/3/3
4	GOL	C	7	-	-	2/4/4/4	-
4	GOL	A	3	-	-	2/4/4/4	-
4	GOL	A	1	-	-	0/4/4/4	-
4	GOL	A	11	-	-	2/4/4/4	-
4	GOL	A	13	-	-	2/4/4/4	-
4	GOL	C	2	-	-	2/4/4/4	-
4	GOL	A	4	-	-	2/4/4/4	-
4	GOL	A	9	-	-	2/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	836	N5P	O4'-C4'	-2.66	1.39	1.45
7	A	836	N5P	CZ2-CE2	-2.41	1.36	1.41
7	A	836	N5P	O3'-C3'	-2.22	1.38	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	836	N5P	PB-O3A-PA	-3.32	121.43	132.83
7	A	836	N5P	CZ2-CE2-NE1	-2.92	129.62	132.11
7	A	836	N5P	PB-O3B-PG	-2.55	124.08	132.83
7	A	836	N5P	O2G-PG-O3B	2.37	112.58	104.64

There are no chirality outliers.

All (27) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	3	GOL	O1-C1-C2-O2
4	A	3	GOL	O1-C1-C2-C3
4	A	4	GOL	O1-C1-C2-C3
4	A	9	GOL	O1-C1-C2-C3
4	A	10	GOL	O1-C1-C2-C3
4	A	11	GOL	O1-C1-C2-C3
4	A	13	GOL	O1-C1-C2-C3
4	C	2	GOL	O1-C1-C2-C3
4	C	5	GOL	O1-C1-C2-C3
4	C	6	GOL	O1-C1-C2-C3
4	C	7	GOL	O1-C1-C2-O2
4	C	7	GOL	O1-C1-C2-C3
4	C	8	GOL	O1-C1-C2-C3
4	A	4	GOL	O1-C1-C2-O2
4	A	9	GOL	O1-C1-C2-O2
4	A	13	GOL	O1-C1-C2-O2
4	C	8	GOL	O1-C1-C2-O2
4	C	6	GOL	C1-C2-C3-O3
4	A	11	GOL	O1-C1-C2-O2
4	C	2	GOL	O1-C1-C2-O2
4	C	5	GOL	O1-C1-C2-O2
4	C	6	GOL	O2-C2-C3-O3
4	A	10	GOL	O1-C1-C2-O2
4	C	6	GOL	O1-C1-C2-O2
7	A	836	N5P	PA-O3A-PB-O2B
4	B	12	GOL	O2-C2-C3-O3
7	A	836	N5P	O4'-C4'-C5'-O5'

There are no ring outliers.

11 monomers are involved in 46 short contacts:

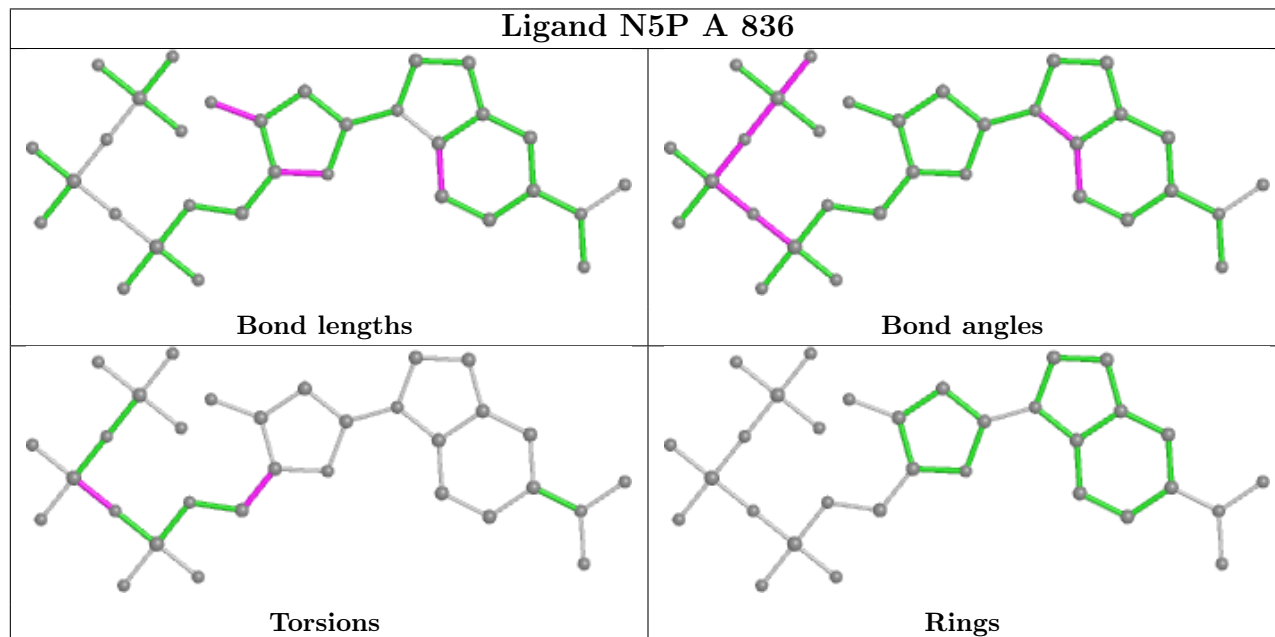
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	6	GOL	3	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	12	GOL	2	0
4	C	5	GOL	2	0
4	A	10	GOL	1	0
8	A	2	FMT	3	0
7	A	836	N5P	21	0
4	A	3	GOL	1	0
4	A	11	GOL	2	0
4	A	13	GOL	6	0
4	C	2	GOL	4	0
4	A	9	GOL	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	533/540 (98%)	0.44	52 (9%) <b>7</b> <b>8</b>	15, 32, 93, 128	0
2	B	11/12 (91%)	-0.53	0 <b>100</b> <b>100</b>	19, 22, 49, 49	0
3	C	12/16 (75%)	-0.44	0 <b>100</b> <b>100</b>	18, 22, 41, 43	0
All	All	556/568 (97%)	0.40	52 (9%) <b>8</b> <b>9</b>	15, 32, 93, 128	0

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	662	ALA	15.0
1	A	645	TRP	13.5
1	A	689	ALA	11.1
1	A	293	ALA	10.7
1	A	643	ALA	10.0
1	A	692	PHE	9.7
1	A	657	LEU	9.7
1	A	669	VAL	9.6
1	A	683	ALA	9.6
1	A	658	MET	9.1
1	A	654	VAL	9.1
1	A	656	PRO	7.3
1	A	668	GLY	7.0
1	A	661	ALA	6.5
1	A	646	MET	6.4
1	A	644	SER	6.1
1	A	655	ASP	6.0
1	A	682	LEU	5.7
1	A	665	ILE	5.6
1	A	693	ILE	5.6
1	A	642	THR	5.0
1	A	664	THR	4.9
1	A	676	HIS	4.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	671	TYR	4.0
1	A	636	ARG	4.0
1	A	681	GLU	3.9
1	A	637	ASP	3.9
1	A	697	PHE	3.9
1	A	684	ILE	3.8
1	A	698	GLN	3.7
1	A	680	GLN	3.5
1	A	640	THR	3.5
1	A	688	GLU	3.4
1	A	639	HIS	3.4
1	A	690	GLN	3.1
1	A	673	MET	3.0
1	A	691	ALA	3.0
1	A	696	TYR	2.8
1	A	666	ASN	2.8
1	A	686	TYR	2.7
1	A	674	SER	2.6
1	A	667	PHE	2.5
1	A	699	SER	2.5
1	A	685	PRO	2.4
1	A	299	TRP	2.4
1	A	663	LYS	2.3
1	A	679	SER	2.3
1	A	672	GLY	2.3
1	A	677	ARG	2.3
1	A	638	ILE	2.3
1	A	503	ILE	2.2
1	A	630	ARG	2.1

## 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	3DR	C	204	11/12	0.94	0.13	41,64,78,79	0
2	DOC	B	112	18/19	0.98	0.10	20,25,31,33	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

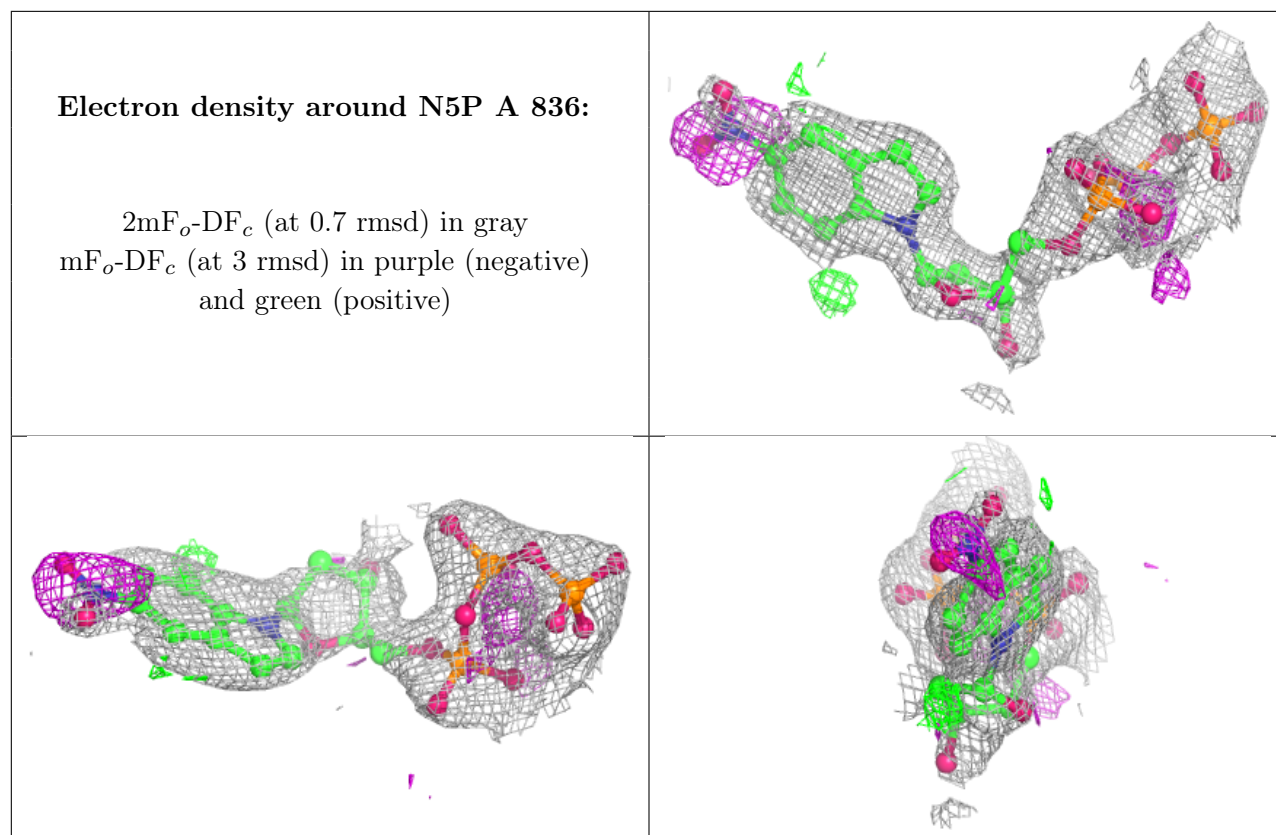
### 6.4 Ligands [i](#)

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
5	MG	A	833	1/1	0.72	0.18	57,57,57,57	1
8	FMT	A	837	3/3	0.75	0.20	57,57,61,63	0
4	GOL	A	10	6/6	0.83	0.24	64,65,66,70	0
5	MG	B	1401	1/1	0.84	0.09	47,47,47,47	0
4	GOL	C	6	6/6	0.84	0.23	43,55,58,63	0
8	FMT	A	838	3/3	0.84	0.13	59,59,65,65	0
7	N5P	A	836	32/32	0.86	0.20	33,45,54,159	32
4	GOL	B	12	6/6	0.87	0.54	64,72,73,74	0
4	GOL	A	3	6/6	0.88	0.15	40,46,56,60	0
4	GOL	A	9	6/6	0.88	0.11	52,57,60,61	0
4	GOL	A	13	6/6	0.89	0.29	62,66,68,73	0
5	MG	C	1501	1/1	0.89	0.04	60,60,60,60	0
4	GOL	C	7	6/6	0.89	0.25	60,64,71,73	0
4	GOL	C	8	6/6	0.89	0.23	55,57,61,64	0
4	GOL	A	11	6/6	0.89	0.17	55,58,63,67	0
4	GOL	C	5	6/6	0.92	0.19	42,53,55,55	0
4	GOL	C	2	6/6	0.92	0.11	37,43,46,47	0
8	FMT	A	839	3/3	0.92	0.11	45,45,54,62	0
5	MG	A	834	1/1	0.94	0.30	34,34,34,34	1
4	GOL	A	4	6/6	0.95	0.17	53,55,61,64	0
6	NA	A	835	1/1	0.95	0.37	77,77,77,77	0
8	FMT	A	2	3/3	0.96	0.21	39,39,39,41	0
4	GOL	A	1	6/6	0.98	0.09	18,21,24,28	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.





## 6.5 Other polymers [i](#)

There are no such residues in this entry.