



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 4, 2023 – 05:54 am GMT

PDB ID : 2UZH  
Title : Mycobacterium smegmatis 2C-methyl-D-erythritol-2,4-cyclodiphosphate synthase (IspF)  
Authors : Buetow, L.; Brown, A.C.; Parish, T.; Hunter, W.N.  
Deposited on : 2007-04-27  
Resolution : 2.20 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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

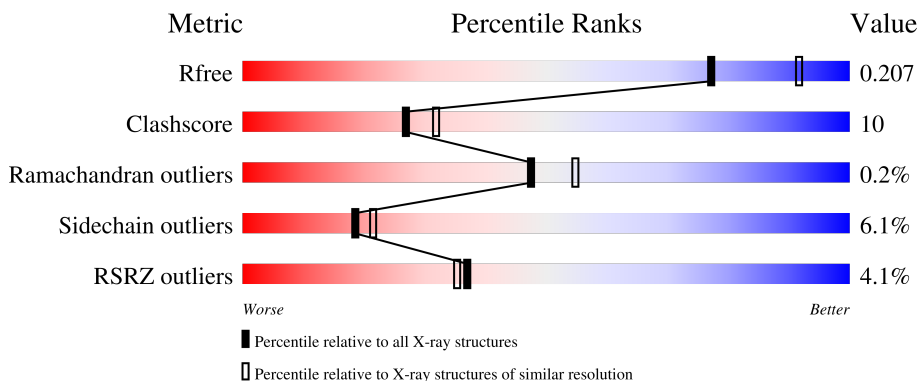
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	165	 5% 78% 15% • 7%
1	B	165	 79% 10% • 7%
1	C	165	 6% 75% 18% • 6%

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
4	EDO	A	1160	-	-	X	-
4	EDO	C	1160	-	-	X	-
6	IPE	B	1160	-	-	-	X
7	PEG	B	1162	-	-	X	-

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 3888 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.

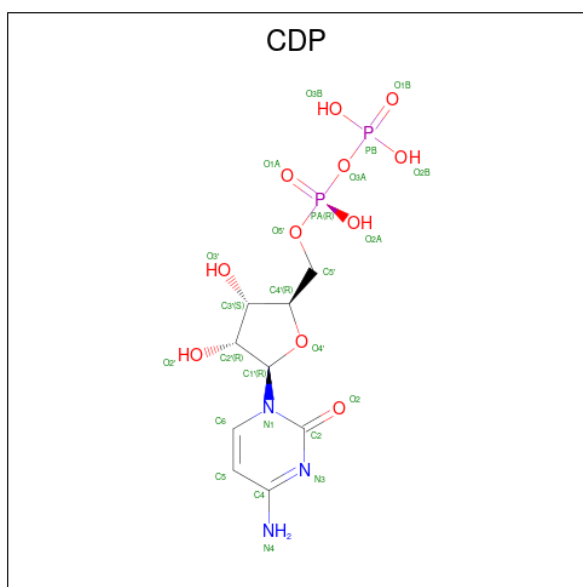
- Molecule 1 is a protein called 2C-METHYL-D-ERYTHRITOL 2,4-CYCLODIPHOSPHATE SYNTHASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	153	1116	695	202	214	5	0	6	0
1	B	153	1140	713	205	217	5	0	11	0
1	C	155	1149	712	209	223	5	0	9	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is CYTIDINE-5'-DIPHOSPHATE (three-letter code: CDP) (formula: C<sub>9</sub>H<sub>15</sub>N<sub>3</sub>O<sub>11</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	Total	C	N	O	P	0	1
			50	18	6	22	4		
3	B	1	Total	C	N	O	P	0	1
			50	18	6	22	4		
3	C	1	Total	C	N	O	P	0	1
			50	18	6	22	4		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	A	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

Continued from previous page...

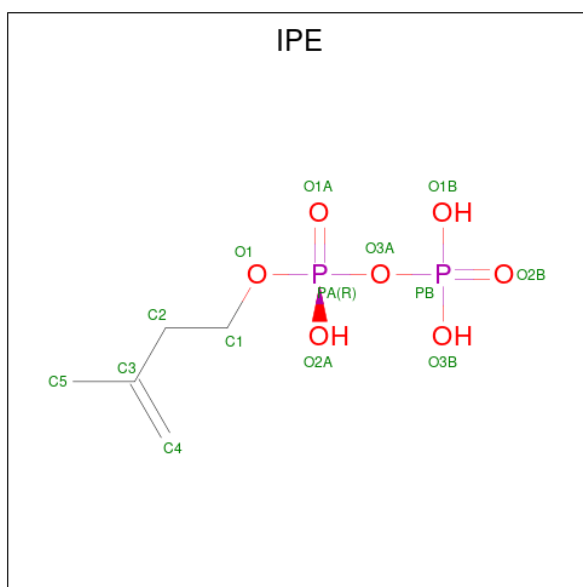
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



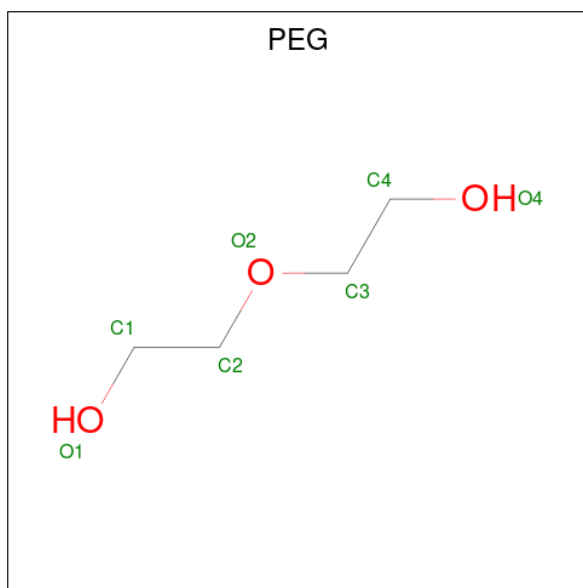
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is 3-METHYLBUT-3-ENYL TRIHYDROGEN DIPHOSPHATE (three-letter code: IPE) (formula:  $C_5H_{12}O_7P_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
6	B	1	14	5	7	2	0	0

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
7	B	1	7	4	3	0	0
7	B	1	7	4	3	0	0

- Molecule 8 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	C	1	Total C O 4 2 2	0	0

- Molecule 9 is water.

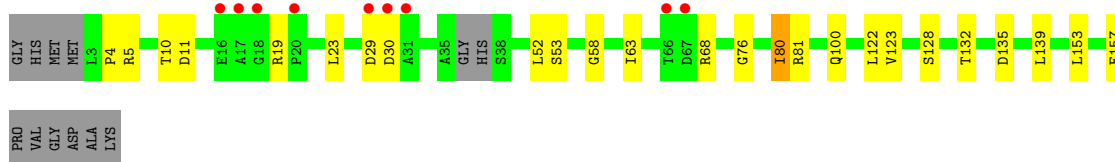
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	80	Total O 80 80	0	0
9	B	109	Total O 109 109	0	0
9	C	83	Total O 83 83	0	0



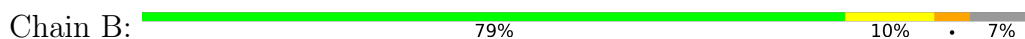
### 3 Residue-property plots [i](#)

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.

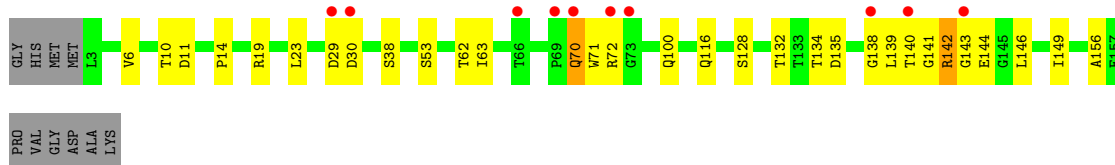
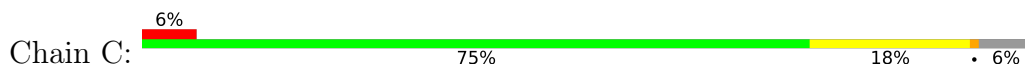
- Molecule 1: 2C-METHYL-D-ERYTHRITOL 2,4-CYCLODIPHOSPHATE SYNTHASE



- Molecule 1: 2C-METHYL-D-ERYTHRITOL 2,4-CYCLODIPHOSPHATE SYNTHASE



- Molecule 1: 2C-METHYL-D-ERYTHRITOL 2,4-CYCLODIPHOSPHATE SYNTHASE



## 4 Data and refinement statistics i

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	159.12Å 159.12Å 54.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	51.30 – 2.20 51.27 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (51.30-2.20) 96.0 (51.27-2.20)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.31 (at 2.20Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.163 , 0.206 0.164 , 0.207	Depositor DCC
$R_{free}$ test set	1671 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.6	Xtrriage
Anisotropy	0.065	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 52.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.019 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3888	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.81% 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: ZN, CDP, PEG, ACT, IPE, EDO, GOL

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	0.56	0/1147	0.71	1/1559 (0.1%)
1	B	0.60	0/1186	0.77	2/1613 (0.1%)
1	C	0.57	0/1191	0.73	0/1619
All	All	0.58	0/3524	0.74	3/4791 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	C	0	1
All	All	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	23	LEU	CA-CB-CG	5.69	128.38	115.30
1	B	142	ARG	NE-CZ-NH2	-5.49	117.56	120.30
1	A	5	ARG	NE-CZ-NH2	-5.16	117.72	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	29	ASP	Peptide
1	C	156	ALA	Peptide

## 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	1116	0	1139	23	0
1	B	1140	0	1182	29	0
1	C	1149	0	1162	26	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	50	0	24	4	0
3	B	50	0	24	7	0
3	C	50	0	24	2	0
4	A	4	0	6	5	0
4	C	4	0	6	11	0
5	A	6	0	8	0	0
5	B	6	0	8	1	0
5	C	6	0	8	1	0
6	B	14	0	9	2	0
7	B	14	0	20	4	0
8	C	4	0	3	1	0
9	A	80	0	0	0	0
9	B	109	0	0	8	0
9	C	83	0	0	0	0
All	All	3888	0	3623	74	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 (74) 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:C:128:SER:HB3	4:C:1160:EDO:H21	1.44	1.00
1:B:140[B]:THR:HG22	9:B:2065:HOH:O	1.71	0.91
1:B:103:GLY:HA2	1:B:140[B]:THR:HG23	1.54	0.90
9:B:2101:HOH:O	1:C:6[A]:VAL:HG13	1.80	0.82
1:A:128:SER:HB3	4:A:1160:EDO:H22	1.64	0.80
4:A:1160:EDO:H21	1:B:58:GLY:H	1.47	0.80
1:B:81[B]:ARG:NH2	7:B:1162:PEG:H42	2.04	0.73

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:81[B]:ARG:HH21	7:B:1162:PEG:H42	1.53	0.72
1:B:102:ILE:HG22	1:B:140[B]:THR:HG21	1.77	0.67
1:B:77:ALA:HB2	7:B:1162:PEG:H11	1.76	0.67
1:C:128:SER:HB3	4:C:1160:EDO:C2	2.25	0.66
3:A:1158[A]:CDP:PB	1:B:11:ASP:OD1	2.54	0.66
1:A:53:SER:HA	4:C:1160:EDO:C2	2.29	0.63
1:A:132:THR:OG1	3:A:1158[A]:CDP:PB	2.58	0.62
1:B:132:THR:OG1	3:B:1158[A]:CDP:PB	2.57	0.62
1:C:139:LEU:HG	1:C:140:THR:HG23	1.82	0.61
4:A:1160:EDO:C2	1:B:58:GLY:H	2.13	0.59
1:C:142:ARG:C	1:C:144[A]:GLU:H	2.06	0.59
1:A:58:GLY:HA2	4:C:1160:EDO:H11	1.83	0.59
1:A:100:GLN:OE1	1:B:10[B]:THR:HG22	2.02	0.59
1:A:76:GLY:O	1:A:80[A]:ILE:HG23	2.03	0.59
1:C:142:ARG:NH2	1:C:144[A]:GLU:OE2	2.35	0.59
9:B:2101:HOH:O	1:C:6[A]:VAL:CG1	2.46	0.58
1:B:140[B]:THR:CG2	9:B:2065:HOH:O	2.37	0.57
1:C:62:THR:HB	1:C:63[A]:ILE:HD12	1.87	0.57
1:A:53:SER:HA	4:C:1160:EDO:H21	1.87	0.55
1:C:142:ARG:C	1:C:144[B]:GLU:H	2.09	0.55
1:A:135:ASP:HB3	1:B:14:PRO:HD3	1.89	0.55
1:B:121:GLU:HG3	9:B:2088:HOH:O	2.06	0.55
1:A:128:SER:HB3	4:A:1160:EDO:C2	2.34	0.54
1:B:77:ALA:CB	7:B:1162:PEG:H11	2.37	0.54
1:A:10:THR:OG1	1:C:100:GLN:NE2	2.41	0.54
1:A:58:GLY:CA	4:C:1160:EDO:H11	2.38	0.54
1:A:80[A]:ILE:HD11	1:A:122:LEU:C	2.29	0.53
1:B:153[B]:LEU:HD12	1:C:6[B]:VAL:HG23	1.91	0.53
1:A:132:THR:HG1	3:A:1158[A]:CDP:PB	2.31	0.53
1:B:135:ASP:HB3	1:C:14:PRO:HD3	1.91	0.52
5:B:1161:GOL:H31	9:B:2108:HOH:O	2.10	0.51
1:B:107:LYS:HA	3:B:1158[B]:CDP:N3	2.26	0.51
3:B:1158[A]:CDP:O3B	1:C:11:ASP:OD1	2.29	0.51
1:C:116:GLN:OE1	4:C:1160:EDO:H12	2.12	0.50
1:A:52:LEU:HD22	1:A:63:ILE:CD1	2.41	0.49
1:B:15:ILE:HD11	1:B:146:LEU:HD13	1.95	0.49
1:B:63[A]:ILE:HD11	1:B:64:PHE:CZ	2.48	0.49
1:C:10:THR:HG22	1:C:149[B]:ILE:HG23	1.94	0.49
1:B:132:THR:HG1	3:B:1158[A]:CDP:PB	2.34	0.49
1:A:80[A]:ILE:HD13	1:A:123:VAL:HG22	1.95	0.48
1:C:134:THR:OG1	1:C:141:GLY:HA2	2.14	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:58:GLY:H	4:C:1160:EDO:C1	2.27	0.47
1:B:142:ARG:NH2	6:B:1160:IPE:O1B	2.46	0.47
3:B:1158[A]:CDP:PB	1:C:11:ASP:OD1	2.72	0.47
1:A:11:ASP:OD1	3:C:1158[B]:CDP:PB	2.73	0.47
5:C:1161:GOL:H32	8:C:1162:ACT:O	2.15	0.45
1:A:4:PRO:HB2	1:A:153[A]:LEU:HD11	1.97	0.45
9:B:2094:HOH:O	1:C:53:SER:HB2	2.16	0.45
1:B:132:THR:OG1	3:B:1158[A]:CDP:O3B	2.34	0.45
1:A:10:THR:HG21	1:C:149[B]:ILE:HD12	1.98	0.44
1:C:116:GLN:OE1	4:C:1160:EDO:C1	2.66	0.44
1:C:29:ASP:O	1:C:30:ASP:HB2	2.17	0.44
1:C:142:ARG:HD2	1:C:143:GLY:N	2.32	0.43
1:C:70[B]:GLN:HG3	1:C:71:TRP:CD1	2.54	0.43
4:A:1160:EDO:H12	1:B:58:GLY:HA2	2.02	0.42
3:B:1158[A]:CDP:O3B	9:B:2106:HOH:O	2.21	0.42
1:C:132:THR:OG1	3:C:1158[B]:CDP:PB	2.78	0.42
1:B:63[A]:ILE:CD1	1:B:82:HIS:HD2	2.34	0.41
6:B:1160:IPE:O3B	1:C:138:GLY:HA2	2.21	0.41
1:A:58:GLY:N	4:C:1160:EDO:H11	2.36	0.41
3:A:1158[A]:CDP:O3B	1:B:11:ASP:OD1	2.38	0.41
1:A:58:GLY:H	4:C:1160:EDO:H11	1.86	0.40
1:A:153[B]:LEU:HD12	1:B:6[B]:VAL:HG13	2.03	0.40
1:B:103:GLY:HA2	1:B:140[B]:THR:CG2	2.39	0.40
1:B:72:ARG:HG3	1:B:72:ARG:NH1	2.37	0.40
1:A:52:LEU:CD2	1:A:63:ILE:HD13	2.52	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	155/165 (94%)	151 (97%)	3 (2%)	1 (1%)	25 26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	160/165 (97%)	160 (100%)	0	0	100	100
1	C	162/165 (98%)	158 (98%)	4 (2%)	0	100	100
All	All	477/495 (96%)	469 (98%)	7 (2%)	1 (0%)	47	55

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	30	ASP

### 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	115/117 (98%)	107 (93%)	8 (7%)	15	16
1	B	120/117 (103%)	114 (95%)	6 (5%)	24	30
1	C	119/117 (102%)	110 (92%)	9 (8%)	13	14
All	All	354/351 (101%)	331 (94%)	23 (6%)	18	19

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	19	ARG
1	A	23	LEU
1	A	68	ARG
1	A	80[A]	ILE
1	A	80[B]	ILE
1	A	81	ARG
1	A	139	LEU
1	A	157	GLU
1	B	23	LEU
1	B	63[A]	ILE
1	B	63[B]	ILE
1	B	72	ARG
1	B	121	GLU

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	146	LEU
1	C	19	ARG
1	C	23	LEU
1	C	38	SER
1	C	70[A]	GLN
1	C	70[B]	GLN
1	C	72	ARG
1	C	135	ASP
1	C	142	ARG
1	C	146	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	82	HIS
1	C	100	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 3 are monoatomic - leaving 15 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
3	CDP	B	1158[A]	-	24,26,26	1.53	3 (12%)	37,40,40	1.29	5 (13%)
5	GOL	A	1161	-	5,5,5	0.35	0	5,5,5	0.43	0
3	CDP	C	1158[A]	-	24,26,26	1.49	5 (20%)	37,40,40	1.06	2 (5%)
3	CDP	B	1158[B]	-	24,26,26	1.34	4 (16%)	37,40,40	1.23	3 (8%)
3	CDP	C	1158[B]	-	24,26,26	1.52	5 (20%)	37,40,40	1.12	3 (8%)
7	PEG	B	1163	-	6,6,6	0.62	0	5,5,5	0.50	0
6	IPE	B	1160	-	11,13,13	2.99	4 (36%)	15,19,19	1.26	2 (13%)
5	GOL	B	1161	-	5,5,5	0.39	0	5,5,5	0.22	0
5	GOL	C	1161	-	5,5,5	0.29	0	5,5,5	0.68	0
3	CDP	A	1158[A]	-	24,26,26	1.37	5 (20%)	37,40,40	1.21	4 (10%)
3	CDP	A	1158[B]	-	24,26,26	1.36	3 (12%)	37,40,40	1.13	3 (8%)
8	ACT	C	1162	-	3,3,3	1.08	0	3,3,3	0.82	0
4	EDO	A	1160	-	3,3,3	0.38	0	2,2,2	0.46	0
7	PEG	B	1162	-	6,6,6	0.65	0	5,5,5	0.42	0
4	EDO	C	1160	-	3,3,3	0.26	0	2,2,2	0.42	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
3	CDP	B	1158[A]	-	-	3/16/32/32	0/2/2/2
5	GOL	A	1161	-	-	0/4/4/4	-
3	CDP	C	1158[A]	-	-	5/16/32/32	0/2/2/2
3	CDP	B	1158[B]	-	-	0/16/32/32	0/2/2/2
3	CDP	C	1158[B]	-	-	5/16/32/32	0/2/2/2
7	PEG	B	1163	-	-	3/4/4/4	-
6	IPE	B	1160	-	-	4/13/13/13	-
5	GOL	B	1161	-	-	2/4/4/4	-
5	GOL	C	1161	-	-	4/4/4/4	-
3	CDP	A	1158[A]	-	-	4/16/32/32	0/2/2/2
3	CDP	A	1158[B]	-	-	6/16/32/32	0/2/2/2
4	EDO	A	1160	-	-	0/1/1/1	-
7	PEG	B	1162	-	-	3/4/4/4	-
4	EDO	C	1160	-	-	0/1/1/1	-

All (29) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	1160	IPE	C4-C3	7.25	1.53	1.33
3	B	1158[A]	CDP	PB-O1B	5.07	1.66	1.50
6	B	1160	IPE	PA-O1A	3.97	1.65	1.50
3	C	1158[B]	CDP	PB-O1B	3.95	1.63	1.50
3	C	1158[A]	CDP	PB-O1B	3.66	1.62	1.50
6	B	1160	IPE	PB-O2B	3.56	1.62	1.50
3	A	1158[B]	CDP	PB-O1B	3.26	1.61	1.50
3	A	1158[A]	CDP	PB-O3B	3.09	1.66	1.54
3	B	1158[B]	CDP	PB-O1B	2.96	1.60	1.50
3	C	1158[A]	CDP	C4-N3	2.82	1.40	1.34
3	A	1158[A]	CDP	C4-N3	2.77	1.40	1.34
3	C	1158[B]	CDP	PB-O3B	2.64	1.65	1.54
3	A	1158[B]	CDP	C4-N3	2.63	1.39	1.34
3	C	1158[B]	CDP	C6-C5	2.51	1.40	1.35
3	B	1158[A]	CDP	C4-N3	2.47	1.39	1.34
3	C	1158[B]	CDP	C4-N3	2.46	1.39	1.34
3	A	1158[A]	CDP	C2-N3	2.46	1.41	1.36
3	B	1158[B]	CDP	C4-N3	2.42	1.39	1.34
3	B	1158[A]	CDP	C6-C5	2.40	1.40	1.35
3	B	1158[B]	CDP	C6-C5	2.33	1.40	1.35
3	A	1158[B]	CDP	C6-C5	2.33	1.40	1.35
3	C	1158[A]	CDP	C2-N3	2.30	1.41	1.36
6	B	1160	IPE	C5-C3	-2.30	1.39	1.48
3	C	1158[A]	CDP	C6-C5	2.29	1.40	1.35
3	C	1158[A]	CDP	O4'-C1'	2.20	1.47	1.42
3	A	1158[A]	CDP	C6-C5	2.17	1.40	1.35
3	B	1158[B]	CDP	C6-N1	2.13	1.43	1.38
3	A	1158[A]	CDP	O4'-C1'	2.10	1.47	1.42
3	C	1158[B]	CDP	C6-N1	2.06	1.43	1.38

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1158[B]	CDP	O2B-PB-O3A	4.08	118.33	104.64
3	A	1158[A]	CDP	PA-O3A-PB	-3.41	121.13	132.83
3	A	1158[B]	CDP	O2B-PB-O3A	3.22	115.42	104.64
3	A	1158[A]	CDP	C2'-C1'-N1	3.08	121.94	113.22
3	B	1158[A]	CDP	O3B-PB-O1B	-3.08	98.63	110.68
6	B	1160	IPE	PA-O3A-PB	-3.05	122.35	132.83
3	C	1158[A]	CDP	O2B-PB-O3A	3.02	114.76	104.64
3	B	1158[A]	CDP	O2B-PB-O3A	2.96	114.58	104.64
3	B	1158[A]	CDP	PA-O3A-PB	-2.87	122.98	132.83
3	C	1158[B]	CDP	PA-O3A-PB	-2.83	123.11	132.83

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1158[B]	CDP	PA-O3A-PB	-2.79	123.27	132.83
3	C	1158[A]	CDP	PA-O3A-PB	-2.46	124.38	132.83
6	B	1160	IPE	O3B-PB-O3A	2.42	112.75	104.64
3	B	1158[B]	CDP	O3B-PB-O1B	-2.40	101.28	110.68
3	C	1158[B]	CDP	O3B-PB-O1B	-2.38	101.36	110.68
3	A	1158[A]	CDP	O2B-PB-O3A	2.33	112.44	104.64
3	B	1158[B]	CDP	PA-O3A-PB	-2.24	125.12	132.83
3	C	1158[B]	CDP	O3B-PB-O3A	2.12	111.74	104.64
3	B	1158[A]	CDP	O4'-C4'-C5'	-2.04	102.66	109.37
3	A	1158[A]	CDP	O4'-C1'-N1	-2.02	103.74	108.36
3	A	1158[B]	CDP	O3B-PB-O1B	-2.01	102.81	110.68
3	B	1158[A]	CDP	O2B-PB-O3B	2.01	115.31	107.64

There are no chirality outliers.

All (39) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1158[A]	CDP	PA-O3A-PB-O2B
3	A	1158[B]	CDP	PB-O3A-PA-O5'
3	A	1158[B]	CDP	C5'-O5'-PA-O1A
3	A	1158[B]	CDP	C5'-O5'-PA-O2A
3	B	1158[A]	CDP	C5'-O5'-PA-O1A
3	B	1158[A]	CDP	C5'-O5'-PA-O2A
3	C	1158[A]	CDP	PA-O3A-PB-O2B
3	C	1158[A]	CDP	C5'-O5'-PA-O1A
3	C	1158[B]	CDP	C5'-O5'-PA-O1A
5	C	1161	GOL	O1-C1-C2-O2
5	C	1161	GOL	O1-C1-C2-C3
5	C	1161	GOL	C1-C2-C3-O3
6	B	1160	IPE	C1-O1-PA-O1A
3	A	1158[B]	CDP	C3'-C4'-C5'-O5'
3	A	1158[B]	CDP	O4'-C4'-C5'-O5'
7	B	1163	PEG	O1-C1-C2-O2
7	B	1163	PEG	O2-C3-C4-O4
5	B	1161	GOL	C1-C2-C3-O3
3	C	1158[A]	CDP	O4'-C4'-C5'-O5'
7	B	1162	PEG	O2-C3-C4-O4
5	B	1161	GOL	O2-C2-C3-O3
6	B	1160	IPE	C1-C2-C3-C5
5	C	1161	GOL	O2-C2-C3-O3
3	A	1158[A]	CDP	C5'-O5'-PA-O3A
3	B	1158[A]	CDP	C5'-O5'-PA-O3A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	C	1158[A]	CDP	C5'-O5'-PA-O3A
3	A	1158[A]	CDP	C5'-O5'-PA-O2A
7	B	1163	PEG	C4-C3-O2-C2
7	B	1162	PEG	C4-C3-O2-C2
3	C	1158[B]	CDP	PB-O3A-PA-O2A
3	C	1158[A]	CDP	C4'-C5'-O5'-PA
7	B	1162	PEG	C1-C2-O2-C3
3	C	1158[B]	CDP	PA-O3A-PB-O1B
3	A	1158[A]	CDP	PA-O3A-PB-O3B
6	B	1160	IPE	PA-O3A-PB-O1B
3	A	1158[B]	CDP	C5'-O5'-PA-O3A
3	C	1158[B]	CDP	C5'-O5'-PA-O3A
6	B	1160	IPE	C1-O1-PA-O3A
3	C	1158[B]	CDP	C3'-C4'-C5'-O5'

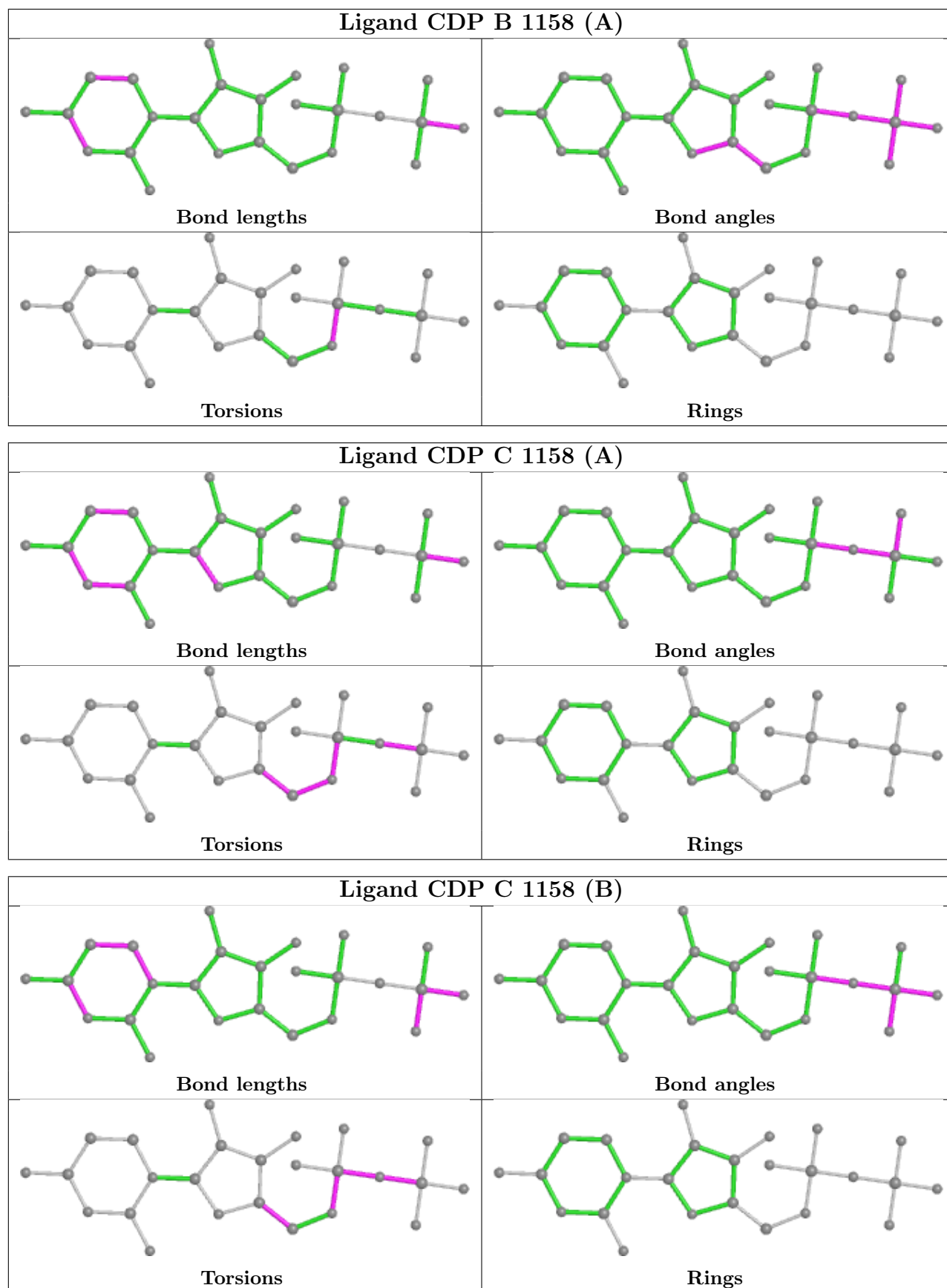
There are no ring outliers.

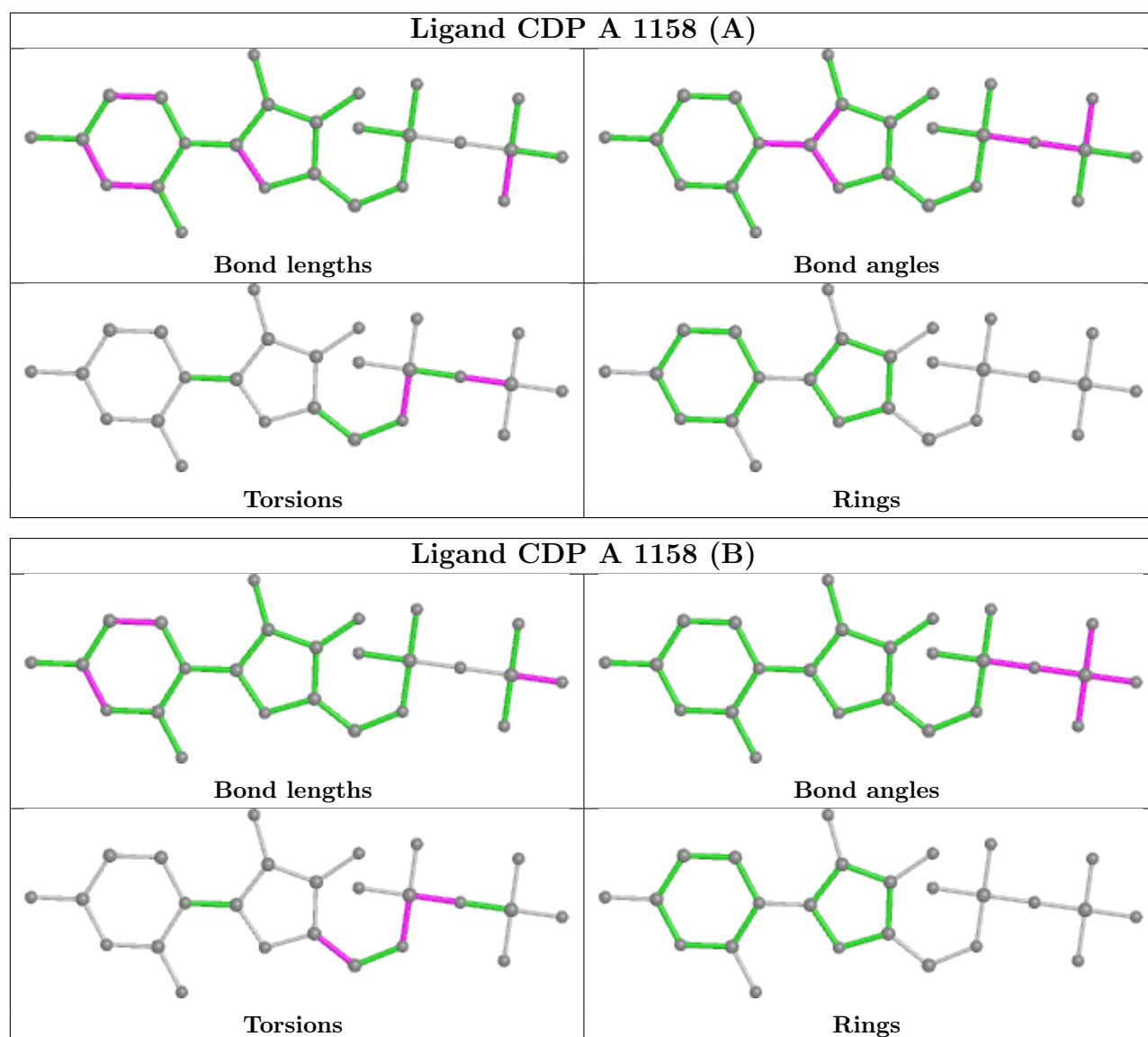
11 monomers are involved in 37 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1158[A]	CDP	6	0
3	B	1158[B]	CDP	1	0
3	C	1158[B]	CDP	2	0
6	B	1160	IPE	2	0
5	B	1161	GOL	1	0
5	C	1161	GOL	1	0
3	A	1158[A]	CDP	4	0
8	C	1162	ACT	1	0
4	A	1160	EDO	5	0
7	B	1162	PEG	4	0
4	C	1160	EDO	11	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	153/165 (92%)	-0.21	9 (5%) 22 21	13, 26, 61, 77	0
1	B	153/165 (92%)	-0.42	0 100 100	9, 18, 37, 54	0
1	C	155/165 (93%)	-0.22	10 (6%) 18 17	12, 25, 64, 73	0
All	All	461/495 (93%)	-0.28	19 (4%) 37 35	9, 23, 57, 77	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	30	ASP	6.7
1	A	29	ASP	5.3
1	A	67	ASP	4.8
1	C	143	GLY	4.8
1	A	66	THR	4.4
1	C	72	ARG	3.9
1	A	17	ALA	3.8
1	C	69	PRO	3.4
1	A	16	GLU	2.8
1	C	66	THR	2.5
1	C	138	GLY	2.5
1	A	20	PRO	2.4
1	C	73	GLY	2.4
1	A	18	GLY	2.4
1	C	29	ASP	2.3
1	A	31	ALA	2.3
1	C	140	THR	2.2
1	C	30	ASP	2.1
1	C	70[A]	GLN	2.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 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.

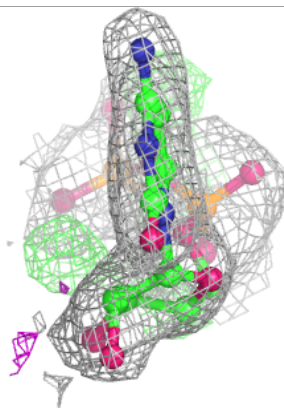
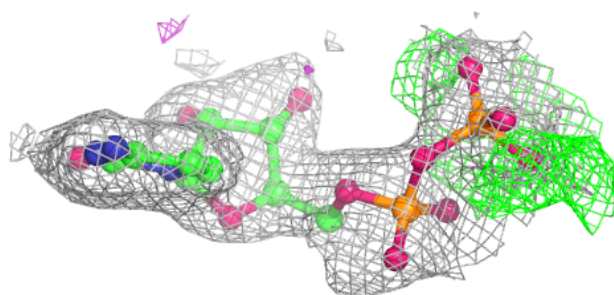
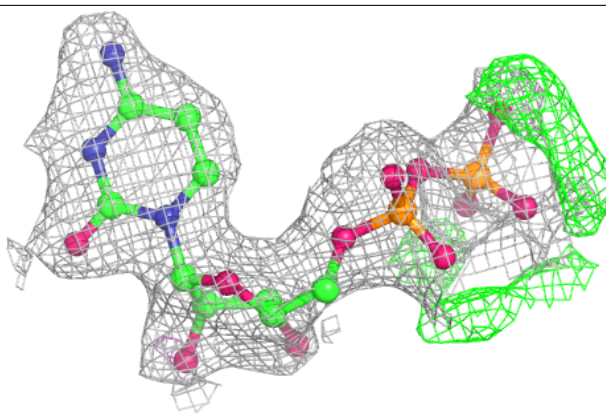
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	IPE	B	1160	14/14	0.62	0.47	44,47,49,50	14
5	GOL	C	1161	6/6	0.77	0.20	49,51,53,53	0
5	GOL	A	1161	6/6	0.80	0.16	60,64,65,65	0
5	GOL	B	1161	6/6	0.84	0.18	57,59,60,60	0
7	PEG	B	1163	7/7	0.84	0.20	41,43,44,45	0
7	PEG	B	1162	7/7	0.86	0.26	38,47,49,50	0
2	ZN	B	1157	1/1	0.91	0.15	39,39,39,39	1
8	ACT	C	1162	4/4	0.93	0.17	37,38,39,39	0
3	CDP	C	1158[B]	25/25	0.94	0.16	19,23,32,34	25
3	CDP	A	1158[A]	25/25	0.94	0.13	9,21,22,23	25
3	CDP	A	1158[B]	25/25	0.94	0.13	16,19,30,30	25
3	CDP	C	1158[A]	25/25	0.94	0.16	44,46,47,48	25
3	CDP	B	1158[B]	25/25	0.95	0.15	14,18,23,24	25
3	CDP	B	1158[A]	25/25	0.95	0.15	6,15,21,25	25
4	EDO	A	1160	4/4	0.97	0.21	19,25,27,27	0
2	ZN	A	1157	1/1	0.98	0.12	41,41,41,41	1
4	EDO	C	1160	4/4	0.99	0.29	30,31,32,34	0
2	ZN	C	1157	1/1	0.99	0.14	26,26,26,26	1

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.

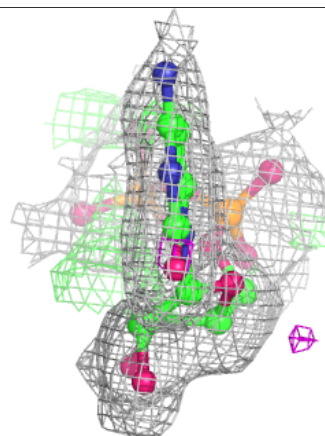
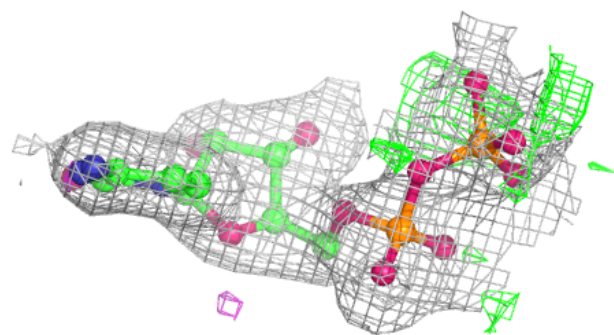
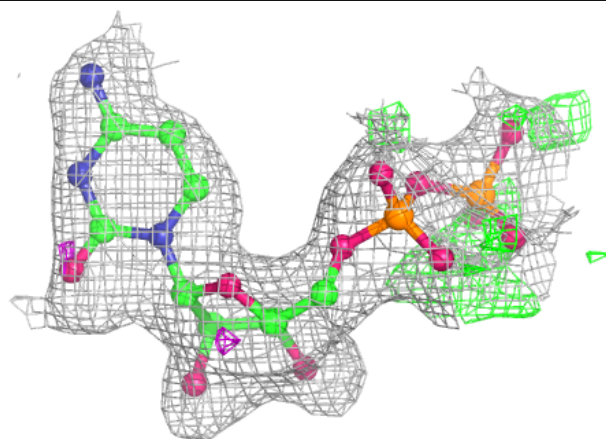


**Electron density around CDP C 1158 (B):**

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

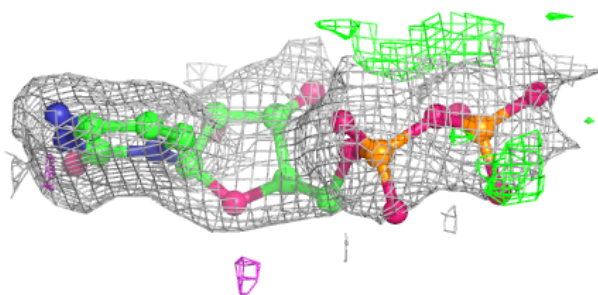
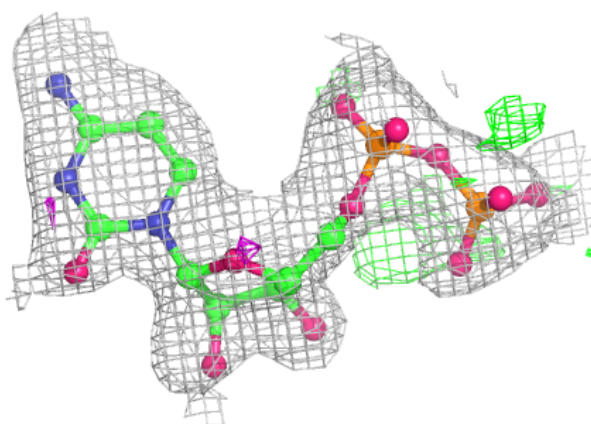
**Electron density around CDP A 1158 (A):**

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

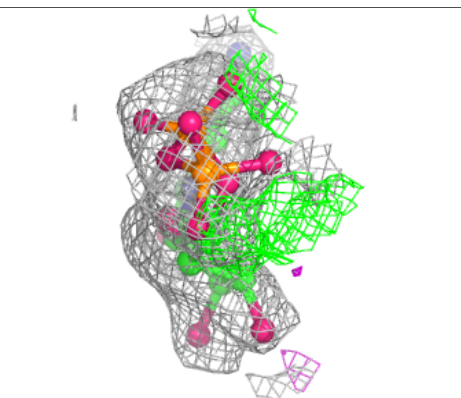
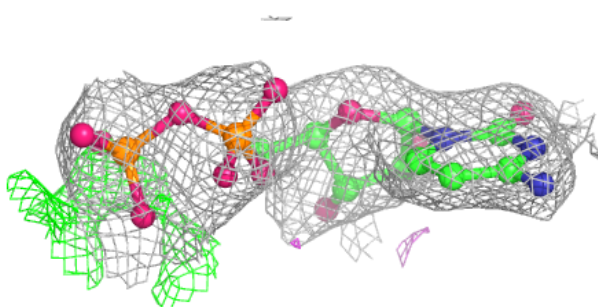
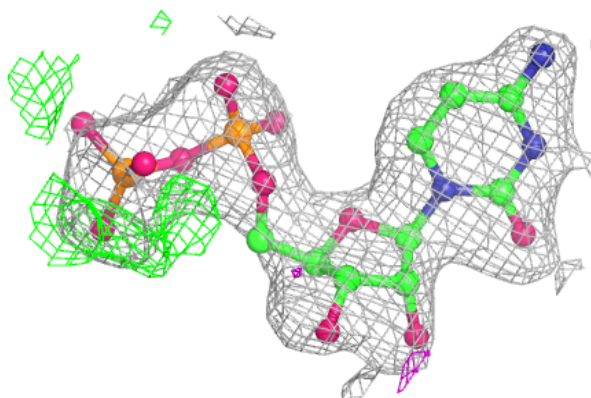


**Electron density around CDP A 1158 (B):**

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

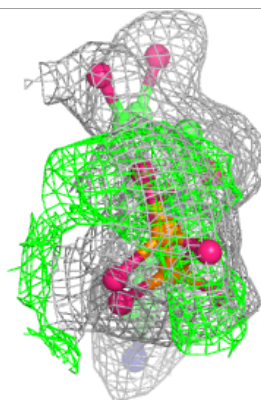
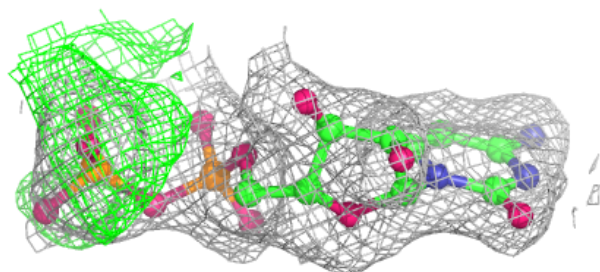
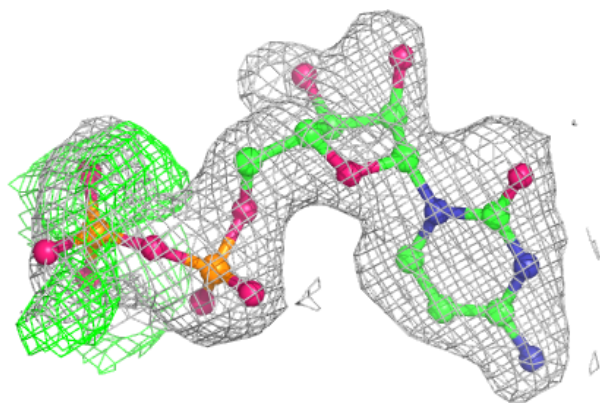
**Electron density around CDP C 1158 (A):**

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

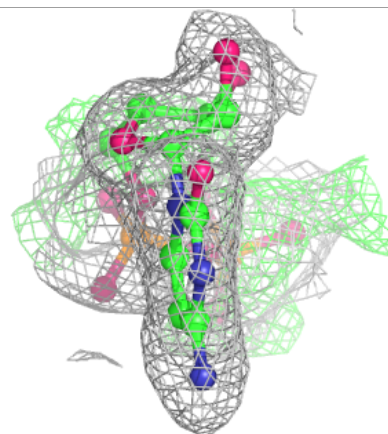
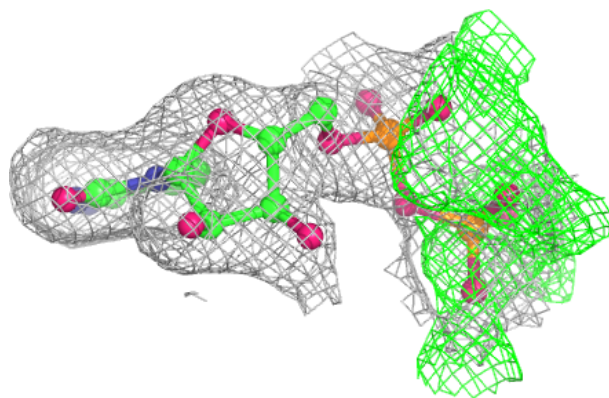
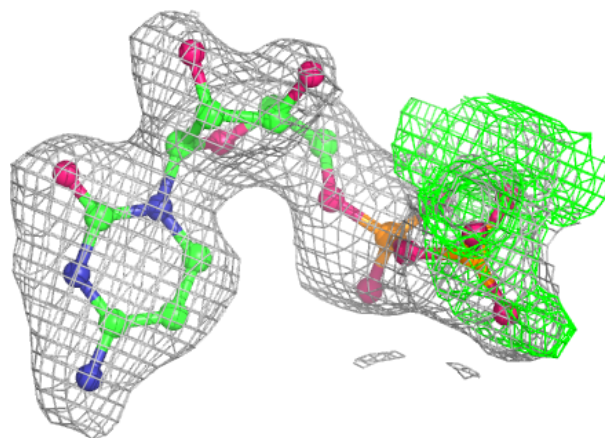


**Electron density around CDP B 1158 (B):**

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

**Electron density around CDP B 1158 (A):**

$2mF_o-DF_c$  (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.