



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 2, 2024 – 04:36 pm GMT

PDB ID : 5AHW  
Title : Crystal structure of universal stress protein MSMEG\_3811 in complex with cAMP  
Authors : Adolph, R.S.; Kleinboelting, S.; Weyand, M.; Steegborn, C.  
Deposited on : 2015-02-10  
Resolution : 2.15 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

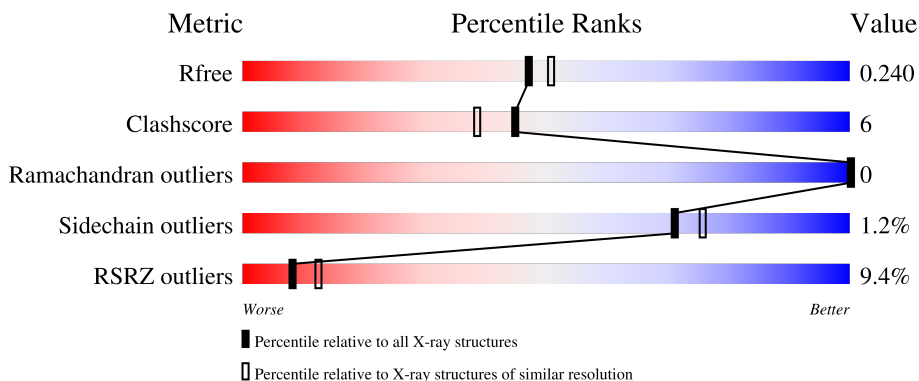
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	147	 8% 72% 12% 15%
1	B	147	 5% 80% 7% 14%
1	C	147	 7% 85% 14% 15%
1	D	147	 7% 81% 5% 14%
1	E	147	 5% 79% 5% 15%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	147	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into four segments: a red segment on the left labeled '18%', a large green segment labeled '76%', a yellow segment labeled '9%', and a grey segment on the far right labeled '15%'. A small black dot is located between the yellow and grey segments.</p>

## 2 Entry composition [i](#)

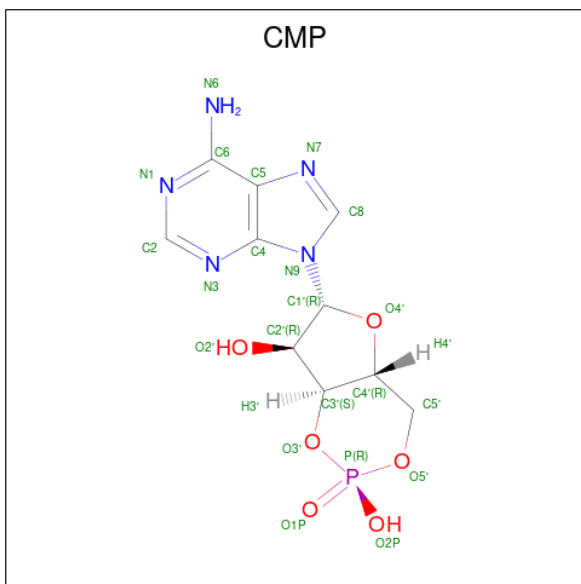
There are 6 unique types of molecules in this entry. The entry contains 6078 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 UNIVERSAL STRESS PROTEIN.

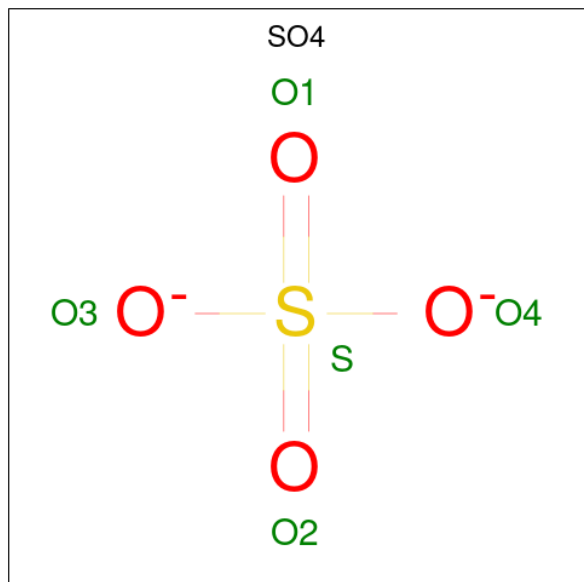
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				
1	A	125	Total	C	N	O	0	1	0	
			900	563	158	179				
1	B	127	Total	C	N	O	0	0	0	
			914	568	164	182				
1	C	146	Total	C	N	O	S	0	3	0
			1066	664	186	215	1			
1	D	126	Total	C	N	O	0	1	0	
			915	571	163	181				
1	E	125	Total	C	N	O	0	0	0	
			902	562	161	179				
1	F	125	Total	C	N	O	0	0	0	
			902	562	161	179				

- Molecule 2 is ADENOSINE-3',5'-CYCLIC-MONOPHOSPHATE (three-letter code: CMP) (formula:  $C_{10}H_{12}N_5O_6P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
2	B	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
2	C	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
2	D	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
2	E	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
2	F	1	Total	C	N	O	P	0	0
			22	10	5	6	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	O S	0	0
			5	4 1		
3	A	1	Total	O S	0	0
			5	4 1		
3	B	1	Total	O S	0	0
			5	4 1		
3	D	1	Total	O S	0	0
			5	4 1		
3	D	1	Total	O S	0	0
			5	4 1		
3	F	1	Total	O S	0	0
			5	4 1		

*Continued on next page...*

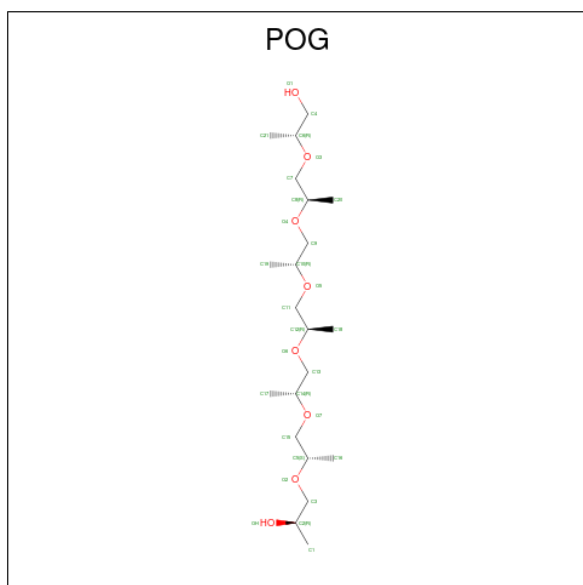
Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	F	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 5 is (20S)-2,5,8,11,14,17-HEXAMETHYL-3,6,9,12,15,18-HEXAOXAHENICOSAN E-1,20-DIOL (three-letter code: POG) (formula: C<sub>21</sub>H<sub>44</sub>O<sub>8</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			29	21	8		
5	D	1	Total	C	O	0	0
			16	12	4		
5	F	1	Total	C	O	0	0
			16	12	4		

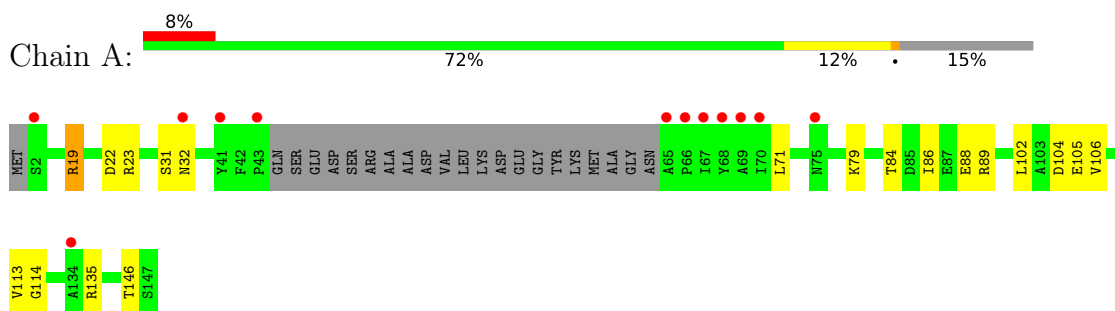
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	41	Total O 41 41	0	0
6	B	52	Total O 52 52	0	0
6	C	56	Total O 56 56	0	0
6	D	42	Total O 42 42	0	0
6	E	37	Total O 37 37	0	0
6	F	21	Total O 21 21	0	0

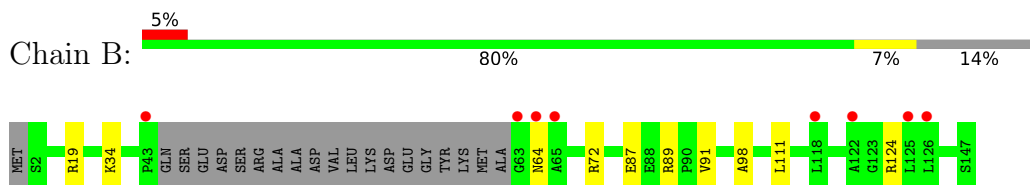
### 3 Residue-property plots [i](#)

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.

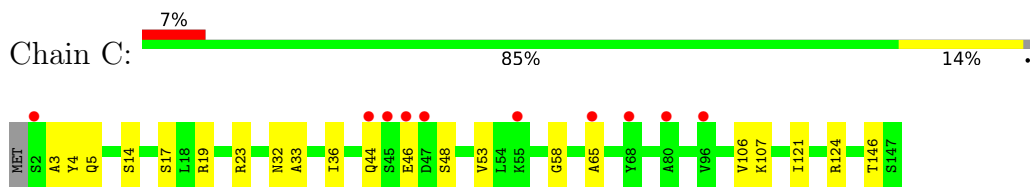
- Molecule 1: UNIVERSAL STRESS PROTEIN



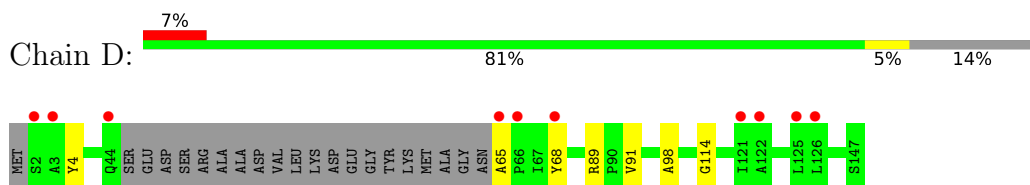
- Molecule 1: UNIVERSAL STRESS PROTEIN



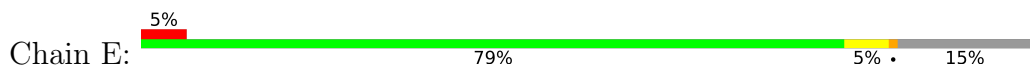
- Molecule 1: UNIVERSAL STRESS PROTEIN



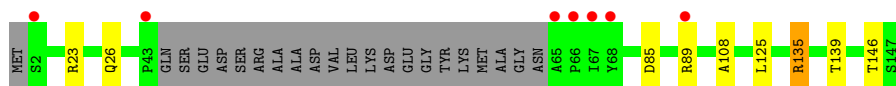
- Molecule 1: UNIVERSAL STRESS PROTEIN



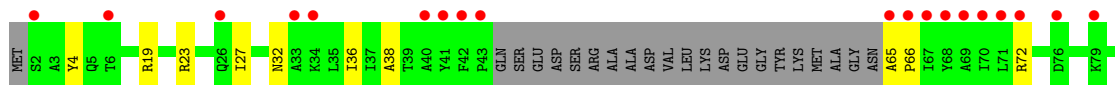
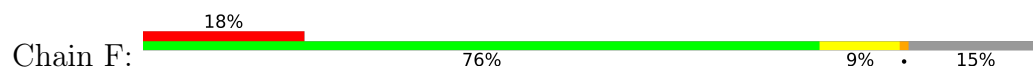
- Molecule 1: UNIVERSAL STRESS PROTEIN







● Molecule 1: UNIVERSAL STRESS PROTEIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.53Å 126.78Å 151.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.64 – 2.15 48.64 – 2.15	Depositor EDS
% Data completeness (in resolution range)	98.8 (48.64-2.15) 98.8 (48.64-2.15)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 2.16Å)	Xtrriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.179 , 0.235 0.187 , 0.240	Depositor DCC
$R_{free}$ test set	2155 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.4	Xtrriage
Anisotropy	0.548	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 47.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6078	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% 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: POG, CMP, SO4, CL

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	0.86	0/911	1.01	6/1242 (0.5%)
1	B	0.92	0/922	0.94	2/1255 (0.2%)
1	C	0.70	0/1085	0.84	0/1477
1	D	0.80	0/926	0.88	0/1261
1	E	0.74	0/910	0.89	2/1239 (0.2%)
1	F	0.63	0/910	0.88	2/1239 (0.2%)
All	All	0.78	0/5664	0.91	12/7713 (0.2%)

There are no bond length outliers.

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	22	ASP	CB-CG-OD1	8.37	125.83	118.30
1	F	124	ARG	NE-CZ-NH1	8.26	124.43	120.30
1	F	124	ARG	NE-CZ-NH2	-8.18	116.21	120.30
1	A	135	ARG	NE-CZ-NH2	-8.11	116.25	120.30
1	A	135	ARG	NE-CZ-NH1	7.44	124.02	120.30
1	A	23	ARG	NE-CZ-NH1	5.84	123.22	120.30
1	E	135	ARG	NE-CZ-NH2	-5.83	117.39	120.30
1	B	89	ARG	NE-CZ-NH2	-5.64	117.48	120.30
1	E	85	ASP	CB-CG-OD1	5.46	123.21	118.30
1	A	19	ARG	NE-CZ-NH2	-5.41	117.59	120.30
1	B	111	LEU	CB-CG-CD2	-5.39	101.84	111.00
1	A	19	ARG	NE-CZ-NH1	5.19	122.90	120.30

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	900	0	925	8	0
1	B	914	0	936	5	0
1	C	1066	0	1090	16	0
1	D	915	0	944	5	0
1	E	902	0	927	7	0
1	F	902	0	927	9	0
2	A	22	0	11	1	0
2	B	22	0	11	1	0
2	C	22	0	11	1	0
2	D	22	0	11	2	0
2	E	22	0	11	2	0
2	F	22	0	11	1	0
3	A	10	0	0	1	0
3	B	5	0	0	0	0
3	D	10	0	0	1	0
3	F	10	0	0	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	C	29	0	44	18	0
5	D	16	0	22	2	0
5	F	16	0	22	4	0
6	A	41	0	0	3	0
6	B	52	0	0	3	1
6	C	56	0	0	19	1
6	D	42	0	0	1	0
6	E	37	0	0	1	0
6	F	21	0	0	1	0
All	All	6078	0	5903	76	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 6.

All (76) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:1148:CMP:H2	2:A:1148:CMP:C2	0.97	1.48
2:C:1148:CMP:H2	2:C:1148:CMP:C2	0.97	1.48
2:F:1148:CMP:H2	2:F:1148:CMP:C2	0.97	1.48
2:B:1148:CMP:H2	2:B:1148:CMP:C2	0.97	1.48
2:D:1148:CMP:C2	2:D:1148:CMP:H2	0.97	1.48
2:E:1148:CMP:H2	2:E:1148:CMP:C2	0.97	1.48
5:C:1149:POG:H172	6:C:2056:HOH:O	1.52	1.07
5:C:1149:POG:H203	6:C:2008:HOH:O	1.55	1.06
5:F:1151:POG:H212	6:F:2007:HOH:O	1.60	1.02
1:B:72:ARG:NH2	6:B:2020:HOH:O	2.01	0.92
1:C:3:ALA:HB1	6:D:2011:HOH:O	1.68	0.92
5:C:1149:POG:C18	6:C:2005:HOH:O	2.16	0.91
5:C:1149:POG:H181	6:C:2050:HOH:O	1.78	0.83
5:C:1149:POG:C8	6:C:2008:HOH:O	2.32	0.78
5:C:1149:POG:C12	6:C:2050:HOH:O	2.33	0.76
5:C:1149:POG:C20	6:C:2008:HOH:O	2.21	0.75
5:C:1149:POG:H181	6:C:2005:HOH:O	1.79	0.73
1:F:124:ARG:HD2	3:F:1149:SO4:O1	1.90	0.71
5:C:1149:POG:O6	6:C:2050:HOH:O	2.07	0.70
1:E:23:ARG:NH1	1:E:26:GLN:OE1	2.24	0.70
1:E:108:ALA:O	1:E:139:THR:HG21	1.91	0.69
1:E:135:ARG:HD2	1:F:124:ARG:HD3	1.76	0.67
5:C:1149:POG:C17	6:C:2056:HOH:O	2.26	0.66
5:C:1149:POG:H152	6:C:2055:HOH:O	1.94	0.66
5:C:1149:POG:C15	6:C:2055:HOH:O	2.44	0.65
1:A:84:THR:OG1	6:A:2013:HOH:O	2.15	0.64
1:A:89:ARG:NH2	1:A:105:GLU:OE2	2.32	0.63
1:C:124:ARG:NH2	6:C:2037:HOH:O	2.05	0.62
5:C:1149:POG:C18	6:C:2050:HOH:O	2.40	0.61
3:D:1149:SO4:O4	5:D:1151:POG:C4	2.49	0.61
3:A:1149:SO4:S	6:A:2034:HOH:O	2.44	0.59
1:F:23:ARG:O	1:F:27:ILE:HD13	2.03	0.57
1:C:36:ILE:HD11	1:C:106[A]:VAL:HG11	1.87	0.57
1:D:65:ALA:O	1:D:68:TYR:CD2	2.57	0.57
1:F:19:ARG:NH1	1:F:146:THR:O	2.37	0.57
1:F:65:ALA:HB1	1:F:66:PRO:HD2	1.86	0.56
1:E:26:GLN:NE2	6:E:2009:HOH:O	2.25	0.56
1:B:19:ARG:NH2	6:B:2010:HOH:O	2.37	0.56
1:F:36:ILE:CD1	1:F:106:VAL:HG11	2.35	0.56
5:C:1149:POG:H183	6:C:2005:HOH:O	1.93	0.55
5:C:1149:POG:C3	6:C:2052:HOH:O	2.56	0.53
1:E:23:ARG:HD2	1:F:4:TYR:CZ	2.44	0.52

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:C:1149:POG:H8	6:C:2008:HOH:O	2.05	0.52
6:B:2026:HOH:O	1:D:89:ARG:NH1	2.42	0.52
1:B:34:LYS:NZ	1:B:87:GLU:OE1	2.38	0.52
1:A:79:LYS:HE3	1:A:86:ILE:HD12	1.92	0.51
6:A:2034:HOH:O	1:B:124:ARG:NH2	2.11	0.51
1:C:121:ILE:HD11	5:F:1151:POG:C4	2.40	0.51
1:B:91:VAL:HG11	1:B:98:ALA:HB1	1.93	0.51
1:C:19:ARG:NH1	1:C:146:THR:O	2.42	0.50
1:C:36:ILE:CD1	1:C:106[A]:VAL:HG11	2.43	0.48
1:A:71:LEU:HD13	1:A:88:GLU:HB3	1.95	0.48
1:A:113[A]:VAL:HG12	1:A:114:GLY:O	2.15	0.47
1:C:58:GLY:HA3	1:C:65:ALA:O	2.14	0.47
1:C:23:ARG:HD2	1:D:4:TYR:CZ	2.51	0.46
1:D:91[A]:VAL:HG12	1:D:98:ALA:HB1	1.97	0.46
1:F:38:ALA:HA	1:F:89:ARG:O	2.16	0.45
1:A:19:ARG:NH1	1:A:146:THR:O	2.45	0.45
1:F:106:VAL:HG12	1:F:106:VAL:O	2.16	0.45
1:C:44:GLN:HG3	1:C:53:VAL:HG13	1.99	0.45
1:A:102:LEU:O	1:A:106:VAL:HG22	2.16	0.44
1:C:4:TYR:O	1:C:33:ALA:HB2	2.17	0.44
5:C:1149:POG:H31	6:C:2052:HOH:O	2.17	0.44
1:D:114:GLY:HA3	2:D:1148:CMP:O1P	2.17	0.43
1:C:14:SER:OG	1:C:17:SER:HB3	2.18	0.43
5:C:1149:POG:C7	6:C:2008:HOH:O	2.61	0.43
1:C:46:GLU:HG2	1:C:46:GLU:O	2.19	0.42
5:F:1151:POG:C11	5:F:1151:POG:H201	2.49	0.42
5:F:1151:POG:C4	5:F:1151:POG:H202	2.49	0.42
5:D:1151:POG:H193	1:E:125:LEU:HD11	2.02	0.41
1:C:5:GLN:O	1:C:33:ALA:HB1	2.21	0.41
1:A:31:SER:O	1:A:32:ASN:HB2	2.20	0.41
1:C:107:LYS:HD3	1:C:107:LYS:HA	1.89	0.41
1:E:146:THR:OG1	2:E:1148:CMP:O1P	2.31	0.41
1:C:5:GLN:C	1:C:33:ALA:HB1	2.41	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)
6:B:2008:HOH:O	6:C:2056:HOH:O[7_545]	1.95	0.25

## 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	122/147 (83%)	120 (98%)	2 (2%)	0	100	100
1	B	123/147 (84%)	123 (100%)	0	0	100	100
1	C	147/147 (100%)	143 (97%)	4 (3%)	0	100	100
1	D	123/147 (84%)	121 (98%)	2 (2%)	0	100	100
1	E	121/147 (82%)	120 (99%)	1 (1%)	0	100	100
1	F	121/147 (82%)	119 (98%)	2 (2%)	0	100	100
All	All	757/882 (86%)	746 (98%)	11 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	93/110 (84%)	92 (99%)	1 (1%)	73	78
1	B	94/110 (86%)	93 (99%)	1 (1%)	73	78
1	C	111/110 (101%)	109 (98%)	2 (2%)	59	63
1	D	95/110 (86%)	95 (100%)	0	100	100
1	E	93/110 (84%)	92 (99%)	1 (1%)	73	78
1	F	93/110 (84%)	91 (98%)	2 (2%)	52	55
All	All	579/660 (88%)	572 (99%)	7 (1%)	71	76

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	104	ASP
1	B	64	ASN
1	C	32	ASN
1	C	48	SER
1	E	89	ARG
1	F	32	ASN
1	F	72	ARG

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

Mol	Chain	Res	Type
1	C	5	GLN
1	C	26	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, 2 are monoatomic - leaving 16 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
2	CMP	D	1148	-	22,25,25	1.73	4 (18%)	24,39,39	1.65	5 (20%)
2	CMP	F	1148	-	22,25,25	1.81	5 (22%)	24,39,39	2.66	9 (37%)
3	SO4	A	1149	-	4,4,4	0.48	0	6,6,6	0.52	0
3	SO4	F	1150	-	4,4,4	0.47	0	6,6,6	0.30	0
5	POG	D	1151	-	12,15,28	0.70	0	14,18,34	1.74	3 (21%)
2	CMP	B	1148	-	22,25,25	1.35	4 (18%)	24,39,39	1.67	6 (25%)
3	SO4	B	1149	-	4,4,4	0.87	0	6,6,6	1.55	2 (33%)
3	SO4	D	1149	-	4,4,4	0.41	0	6,6,6	0.77	0
3	SO4	D	1150	-	4,4,4	0.30	0	6,6,6	0.57	0
3	SO4	F	1149	-	4,4,4	0.50	0	6,6,6	0.69	0
5	POG	C	1149	-	21,28,28	1.25	3 (14%)	26,34,34	2.37	10 (38%)
5	POG	F	1151	-	12,15,28	1.13	2 (16%)	14,18,34	1.99	4 (28%)
2	CMP	E	1148	-	22,25,25	1.79	4 (18%)	24,39,39	2.47	8 (33%)
2	CMP	A	1148	-	22,25,25	1.73	6 (27%)	24,39,39	2.86	10 (41%)
3	SO4	A	1150	-	4,4,4	0.66	0	6,6,6	0.74	0
2	CMP	C	1148	-	22,25,25	1.36	3 (13%)	24,39,39	1.78	6 (25%)

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
2	CMP	D	1148	-	-	0/0/31/31	0/4/4/4
2	CMP	F	1148	-	-	0/0/31/31	0/4/4/4
5	POG	D	1151	-	-	9/15/15/32	-
2	CMP	B	1148	-	-	0/0/31/31	0/4/4/4
5	POG	C	1149	-	-	17/32/32/32	-
5	POG	F	1151	-	-	8/15/15/32	-
2	CMP	E	1148	-	-	0/0/31/31	0/4/4/4
2	CMP	A	1148	-	-	0/0/31/31	0/4/4/4
2	CMP	C	1148	-	-	0/0/31/31	0/4/4/4

All (31) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	1148	CMP	P-O5'	5.31	1.63	1.57
2	F	1148	CMP	P-O5'	4.67	1.63	1.57
2	D	1148	CMP	P-O5'	4.49	1.62	1.57

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	1148	CMP	P-O3'	4.03	1.64	1.57
2	A	1148	CMP	O5'-C5'	-3.97	1.40	1.46
2	F	1148	CMP	O5'-C5'	-3.79	1.40	1.46
2	D	1148	CMP	P-O3'	3.70	1.63	1.57
2	B	1148	CMP	P-O3'	3.13	1.63	1.57
2	A	1148	CMP	P-O5'	3.12	1.61	1.57
5	C	1149	POG	O5-C11	3.09	1.49	1.43
2	C	1148	CMP	O5'-C5'	-3.09	1.41	1.46
2	C	1148	CMP	P-O5'	2.91	1.61	1.57
2	C	1148	CMP	O3'-C3'	-2.80	1.40	1.44
2	F	1148	CMP	O3'-C3'	-2.72	1.40	1.44
5	C	1149	POG	C13-C14	2.70	1.56	1.50
5	C	1149	POG	C15-C5	2.54	1.56	1.50
2	A	1148	CMP	P-O3'	2.46	1.61	1.57
2	D	1148	CMP	O5'-C5'	-2.44	1.42	1.46
2	F	1148	CMP	C5-C4	2.43	1.47	1.40
2	F	1148	CMP	O4'-C4'	-2.28	1.39	1.45
2	E	1148	CMP	C5-C4	2.27	1.46	1.40
2	A	1148	CMP	O3'-C3'	-2.26	1.41	1.44
2	A	1148	CMP	O4'-C4'	-2.23	1.40	1.45
2	B	1148	CMP	C5-N7	-2.22	1.31	1.39
5	F	1151	POG	O4-C9	2.17	1.47	1.43
2	E	1148	CMP	O5'-C5'	-2.16	1.43	1.46
5	F	1151	POG	C9-C10	2.13	1.55	1.50
2	A	1148	CMP	O4'-C1'	2.13	1.44	1.41
2	B	1148	CMP	C2-N3	2.08	1.35	1.32
2	B	1148	CMP	C5-C4	2.06	1.46	1.40
2	D	1148	CMP	C5-N7	-2.05	1.32	1.39

All (63) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	1148	CMP	O3'-C3'-C4'	-9.28	103.70	110.71
2	A	1148	CMP	O3'-C3'-C4'	-7.84	104.79	110.71
2	E	1148	CMP	O2P-P-O1P	6.19	128.13	108.73
5	C	1149	POG	C13-O6-C12	5.80	123.67	115.02
2	A	1148	CMP	O3'-C3'-C2'	5.46	120.95	115.61
2	E	1148	CMP	O5'-P-O3'	-5.43	98.20	105.68
5	D	1151	POG	C11-O5-C10	5.10	122.63	115.02
2	E	1148	CMP	O3'-C3'-C2'	5.01	120.52	115.61
5	F	1151	POG	O4-C9-C10	4.89	122.70	110.90
2	A	1148	CMP	O2P-P-O1P	4.61	123.16	108.73

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1148	CMP	N3-C2-N1	-4.48	121.67	128.68
5	C	1149	POG	O6-C12-C11	4.40	122.93	108.89
2	F	1148	CMP	O2P-P-O1P	4.31	122.22	108.73
2	B	1148	CMP	N3-C2-N1	-4.22	122.08	128.68
2	D	1148	CMP	N3-C2-N1	-4.17	122.16	128.68
2	A	1148	CMP	N3-C2-N1	-4.10	122.27	128.68
5	C	1149	POG	C3-O2-C5	4.07	121.09	115.02
2	A	1148	CMP	O2P-P-O3'	3.92	116.18	107.04
2	A	1148	CMP	C4-C5-N7	-3.82	105.41	109.40
2	F	1148	CMP	O3'-C3'-C2'	3.74	119.27	115.61
2	C	1148	CMP	O2P-P-O1P	3.57	119.91	108.73
2	E	1148	CMP	O3'-C3'-C4'	-3.46	108.09	110.71
2	E	1148	CMP	O5'-P-O1P	-3.43	102.53	110.44
5	C	1149	POG	C15-O7-C14	-3.42	109.92	115.02
5	C	1149	POG	O7-C14-C13	3.41	119.76	108.89
5	C	1149	POG	C18-C12-C11	-3.37	102.96	112.63
2	B	1148	CMP	O3'-C3'-C4'	-3.28	108.23	110.71
2	F	1148	CMP	N3-C2-N1	-3.16	123.74	128.68
5	F	1151	POG	O5-C11-C12	3.12	114.90	108.86
2	A	1148	CMP	O5'-P-O3'	-2.96	101.60	105.68
2	D	1148	CMP	O3'-C3'-C4'	-2.95	108.48	110.71
5	C	1149	POG	C19-C10-C9	-2.84	104.46	112.63
2	E	1148	CMP	C4-C5-N7	-2.83	106.45	109.40
2	F	1148	CMP	O5'-P-O3'	-2.81	101.81	105.68
2	D	1148	CMP	O2P-P-O1P	2.75	117.34	108.73
2	B	1148	CMP	O5'-P-O3'	-2.72	101.92	105.68
3	B	1149	SO4	O4-S-O3	2.68	120.50	109.06
2	A	1148	CMP	O2P-P-O5'	-2.67	100.65	107.16
5	C	1149	POG	C7-O3-C6	-2.65	111.07	115.02
5	D	1151	POG	O4-C8-C7	2.59	117.14	108.89
2	C	1148	CMP	O5'-P-O3'	-2.52	102.21	105.68
2	C	1148	CMP	O5'-C5'-C4'	-2.52	99.83	105.71
2	B	1148	CMP	O2P-P-O5'	2.51	113.28	107.16
2	C	1148	CMP	C4-C5-N7	-2.47	106.83	109.40
2	D	1148	CMP	C2-N1-C6	2.47	122.97	118.75
3	B	1149	SO4	O4-S-O2	-2.41	96.74	109.31
5	C	1149	POG	C17-C14-C13	-2.32	105.96	112.63
2	B	1148	CMP	C2-N1-C6	2.32	122.71	118.75
2	F	1148	CMP	C2-N1-C6	2.31	122.71	118.75
5	F	1151	POG	C9-O4-C8	-2.30	111.59	115.02
5	F	1151	POG	O5-C10-C9	2.28	116.14	108.89
2	F	1148	CMP	O5'-P-O1P	-2.26	105.22	110.44

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1148	CMP	O5'-P-O3'	-2.26	102.57	105.68
2	B	1148	CMP	O3'-C3'-C2'	2.24	117.80	115.61
2	F	1148	CMP	N6-C6-N1	2.18	123.11	118.57
5	D	1151	POG	C20-C8-C7	-2.18	106.36	112.63
2	F	1148	CMP	O2P-P-O5'	2.17	112.46	107.16
2	E	1148	CMP	N3-C2-N1	-2.15	125.32	128.68
2	A	1148	CMP	O3'-P-O1P	-2.14	105.81	110.39
2	E	1148	CMP	C2'-C3'-C4'	-2.12	99.47	103.22
5	C	1149	POG	O4-C8-C7	-2.07	102.29	108.89
2	A	1148	CMP	C1'-N9-C4	-2.06	123.03	126.64
2	C	1148	CMP	C1'-N9-C4	-2.06	123.03	126.64

There are no chirality outliers.

All (34) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	1149	POG	C16-C5-O2-C3
5	C	1149	POG	O7-C15-C5-O2
5	C	1149	POG	O7-C15-C5-C16
5	C	1149	POG	O1-C4-C6-O3
5	C	1149	POG	O1-C4-C6-C21
5	C	1149	POG	O3-C7-C8-O4
5	C	1149	POG	O5-C11-C12-O6
5	C	1149	POG	O5-C11-C12-C18
5	D	1151	POG	O3-C7-C8-O4
5	D	1151	POG	O3-C7-C8-C20
5	D	1151	POG	C19-C10-C9-O4
5	D	1151	POG	O5-C11-C12-C18
5	F	1151	POG	O3-C7-C8-O4
5	F	1151	POG	O3-C7-C8-C20
5	F	1151	POG	C19-C10-C9-O4
5	C	1149	POG	C10-C9-O4-C8
5	C	1149	POG	C12-C11-O5-C10
5	D	1151	POG	C8-C7-O3-C6
5	D	1151	POG	C10-C9-O4-C8
5	F	1151	POG	C10-C9-O4-C8
5	F	1151	POG	C12-C11-O5-C10
5	C	1149	POG	O3-C7-C8-C20
5	D	1151	POG	O5-C10-C9-O4
5	F	1151	POG	O5-C10-C9-O4
5	F	1151	POG	C4-C6-O3-C7
5	C	1149	POG	C14-C13-O6-C12

*Continued on next page...*

*Continued from previous page...*

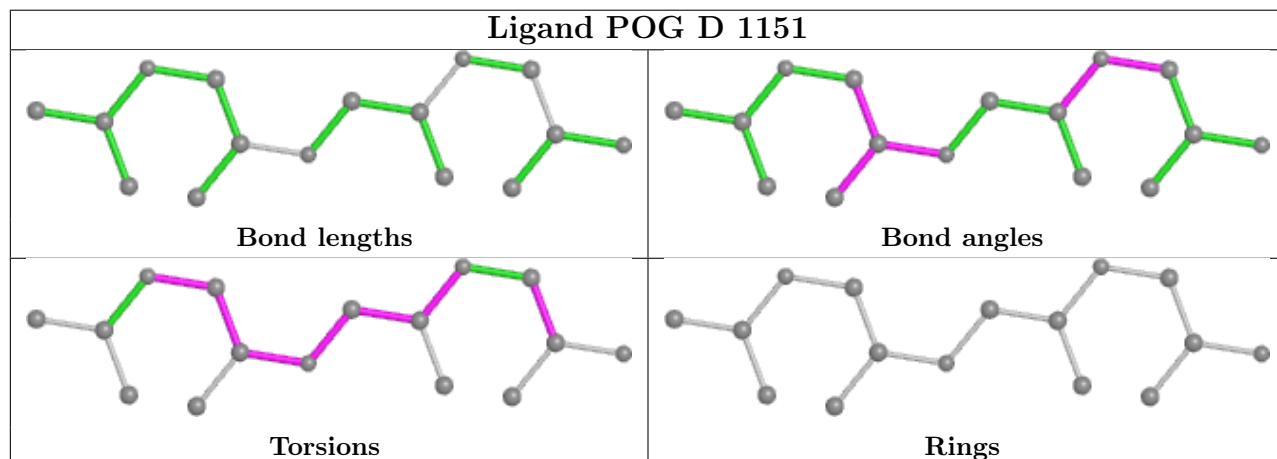
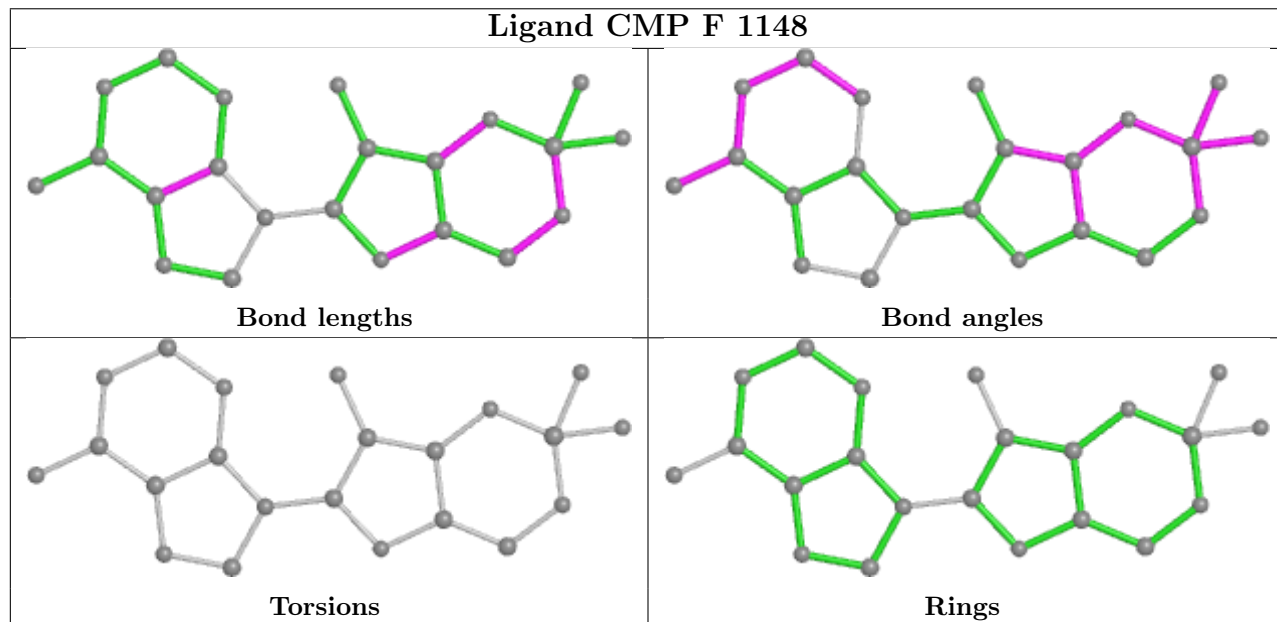
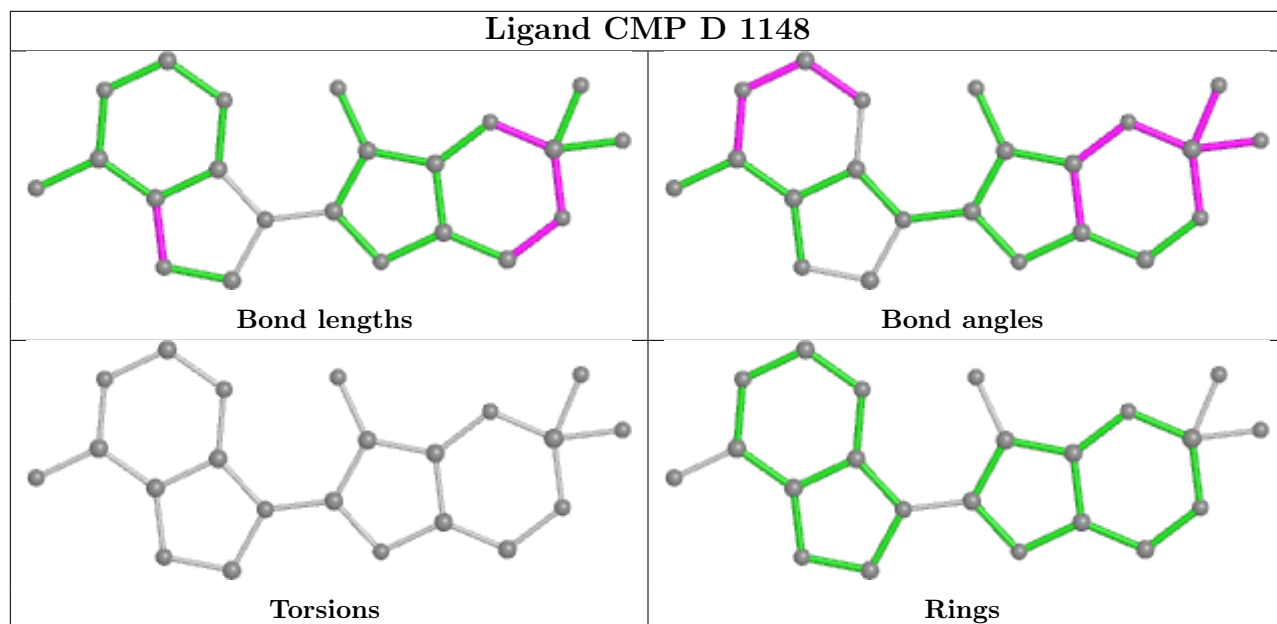
Mol	Chain	Res	Type	Atoms
5	C	1149	POG	C5-C15-O7-C14
5	F	1151	POG	C21-C6-O3-C7
5	C	1149	POG	O6-C13-C14-O7
5	C	1149	POG	C21-C6-O3-C7
5	C	1149	POG	C20-C8-O4-C9
5	C	1149	POG	C19-C10-O5-C11
5	D	1151	POG	C20-C8-O4-C9
5	D	1151	POG	C9-C10-O5-C11

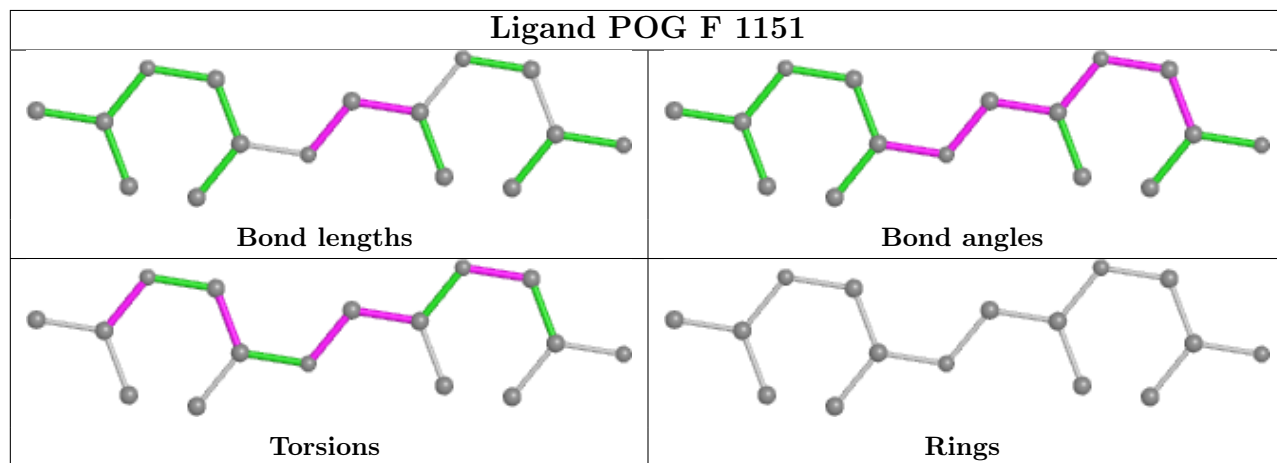
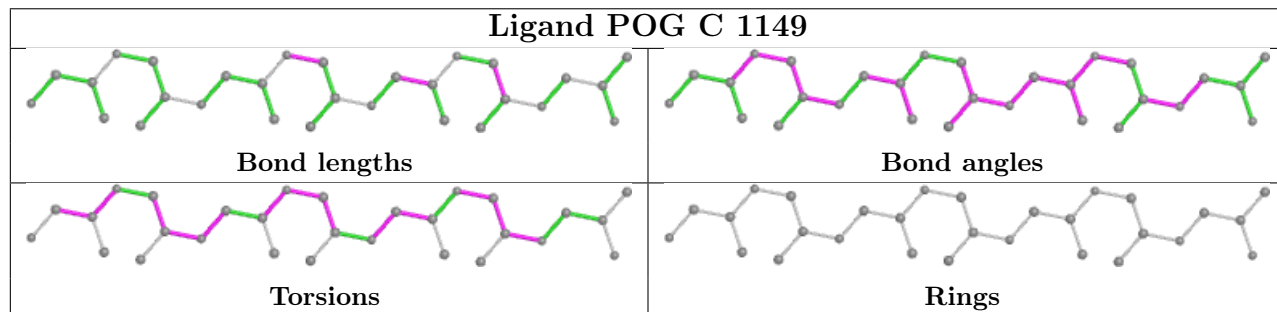
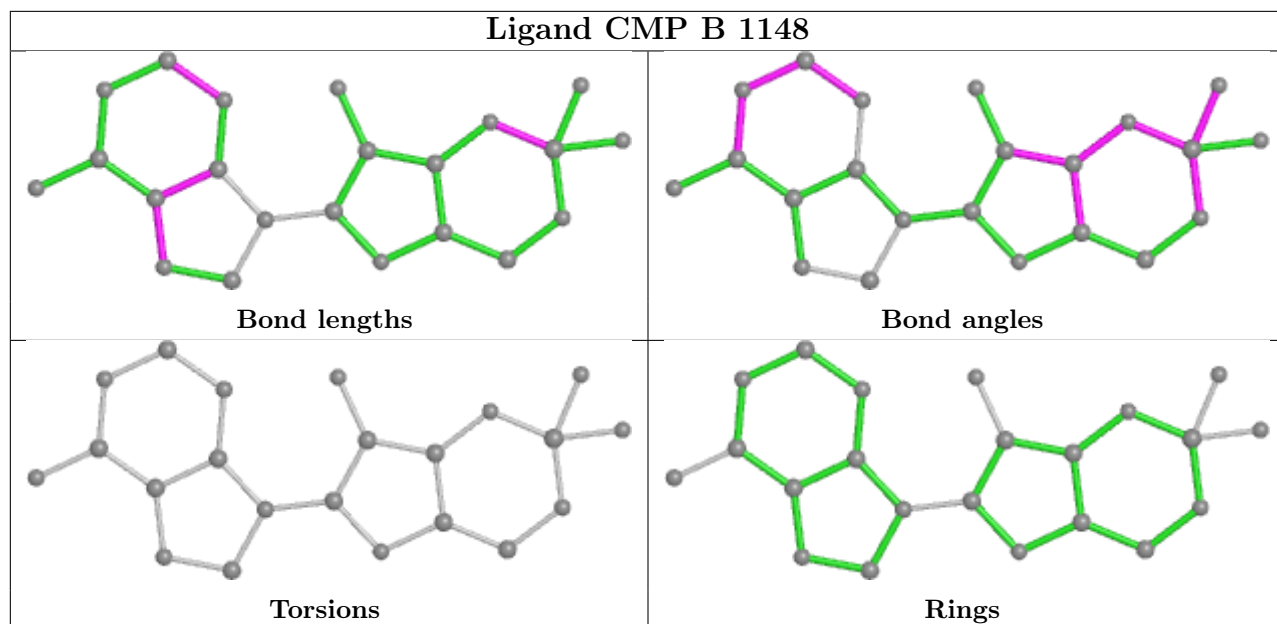
There are no ring outliers.

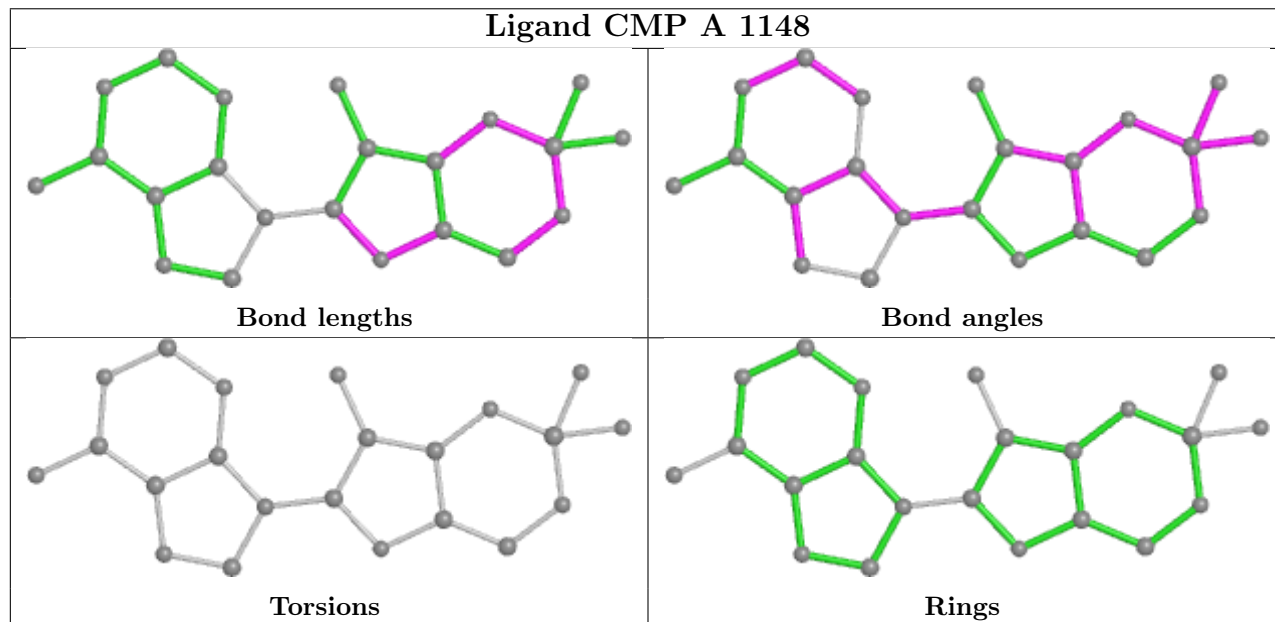
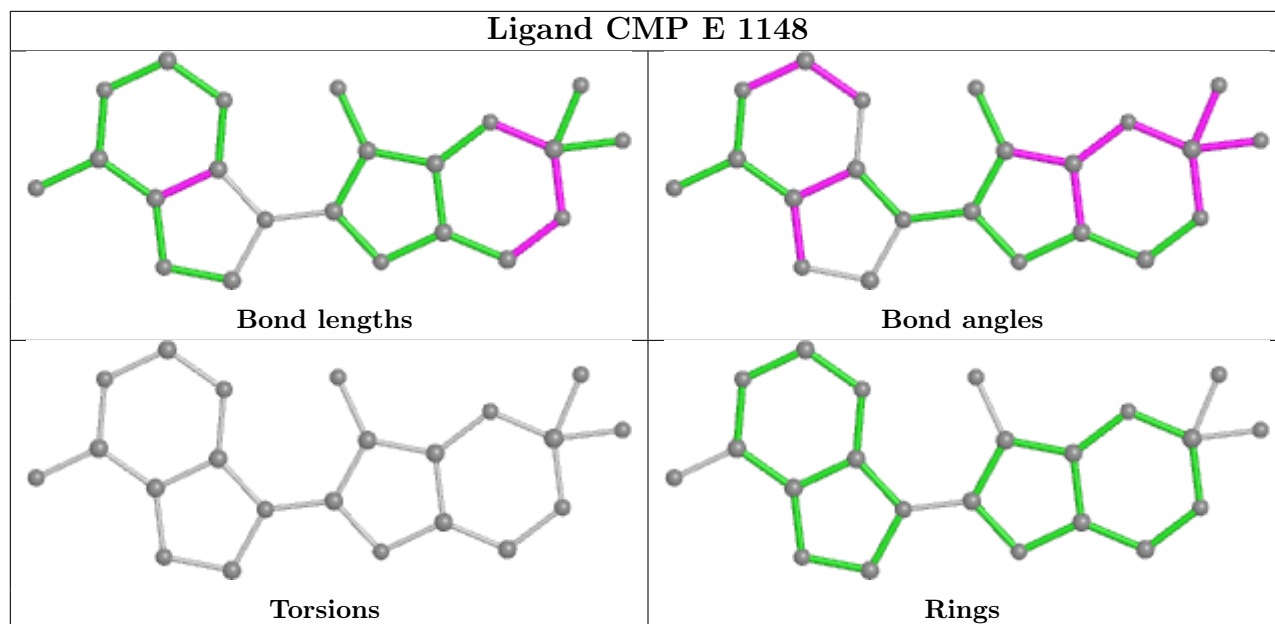
12 monomers are involved in 34 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1148	CMP	2	0
2	F	1148	CMP	1	0
3	A	1149	SO4	1	0
5	D	1151	POG	2	0
2	B	1148	CMP	1	0
3	D	1149	SO4	1	0
3	F	1149	SO4	1	0
5	C	1149	POG	18	0
5	F	1151	POG	4	0
2	E	1148	CMP	2	0
2	A	1148	CMP	1	0
2	C	1148	CMP	1	0

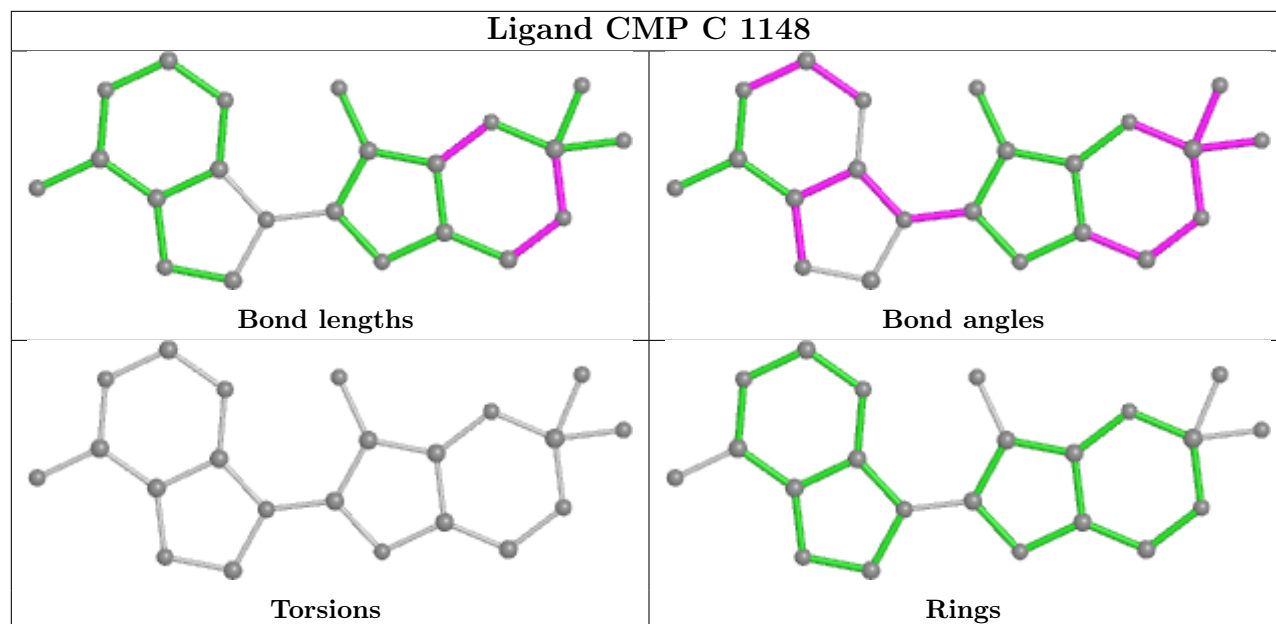
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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	125/147 (85%)	0.37	12 (9%) 8 11	25, 44, 72, 96	0
1	B	127/147 (86%)	0.25	8 (6%) 20 27	23, 36, 58, 83	0
1	C	146/147 (99%)	0.37	10 (6%) 17 24	27, 51, 79, 116	0
1	D	126/147 (85%)	0.29	10 (7%) 12 17	27, 40, 67, 93	0
1	E	125/147 (85%)	0.17	7 (5%) 24 33	31, 49, 77, 109	0
1	F	125/147 (85%)	0.89	26 (20%) 1 1	32, 62, 94, 107	0
All	All	774/882 (87%)	0.39	73 (9%) 8 12	23, 47, 85, 116	0

All (73) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	86	ILE	8.0
1	D	65	ALA	7.2
1	F	68	TYR	6.4
1	F	67	ILE	6.4
1	C	2	SER	5.7
1	F	66	PRO	5.5
1	F	42	PHE	5.4
1	C	45	SER	5.3
1	F	33	ALA	5.2
1	A	68	TYR	4.9
1	A	2	SER	4.6
1	E	68	TYR	4.6
1	F	43	PRO	4.5
1	A	69	ALA	4.4
1	C	47	ASP	4.2
1	D	121	ILE	3.9
1	F	41	TYR	3.6
1	B	125	LEU	3.6
1	F	90	PRO	3.6

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	43	PRO	3.4
1	D	3	ALA	3.4
1	E	67	ILE	3.4
1	B	65	ALA	3.3
1	F	65	ALA	3.2
1	E	2	SER	3.2
1	F	69	ALA	3.2
1	E	43	PRO	3.2
1	F	71	LEU	3.2
1	B	64	ASN	3.1
1	F	79	LYS	3.1
1	B	63	GLY	3.0
1	F	72	ARG	2.9
1	A	65	ALA	2.9
1	F	76	ASP	2.9
1	C	55	LYS	2.8
1	C	44	GLN	2.8
1	A	41	TYR	2.8
1	E	65	ALA	2.7
1	B	126	LEU	2.7
1	F	121	ILE	2.7
1	C	46	GLU	2.7
1	E	89	ARG	2.7
1	F	2	SER	2.6
1	E	66	PRO	2.6
1	D	44	GLN	2.6
1	F	6	THR	2.6
1	F	34	LYS	2.6
1	A	66	PRO	2.6
1	A	70	ILE	2.5
1	F	89	ARG	2.5
1	D	125	LEU	2.5
1	A	67	ILE	2.5
1	B	118	LEU	2.5
1	D	68	TYR	2.5
1	C	80	ALA	2.4
1	C	65	ALA	2.4
1	D	126	LEU	2.4
1	D	2	SER	2.4
1	D	122	ALA	2.3
1	F	70	ILE	2.3
1	F	88	GLU	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	125	LEU	2.2
1	A	75	ASN	2.2
1	B	122	ALA	2.2
1	B	43	PRO	2.2
1	A	32	ASN	2.1
1	C	96	VAL	2.1
1	A	134	ALA	2.1
1	C	68	TYR	2.1
1	F	122	ALA	2.1
1	D	66	PRO	2.0
1	F	40	ALA	2.0
1	F	26	GLN	2.0

## 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.

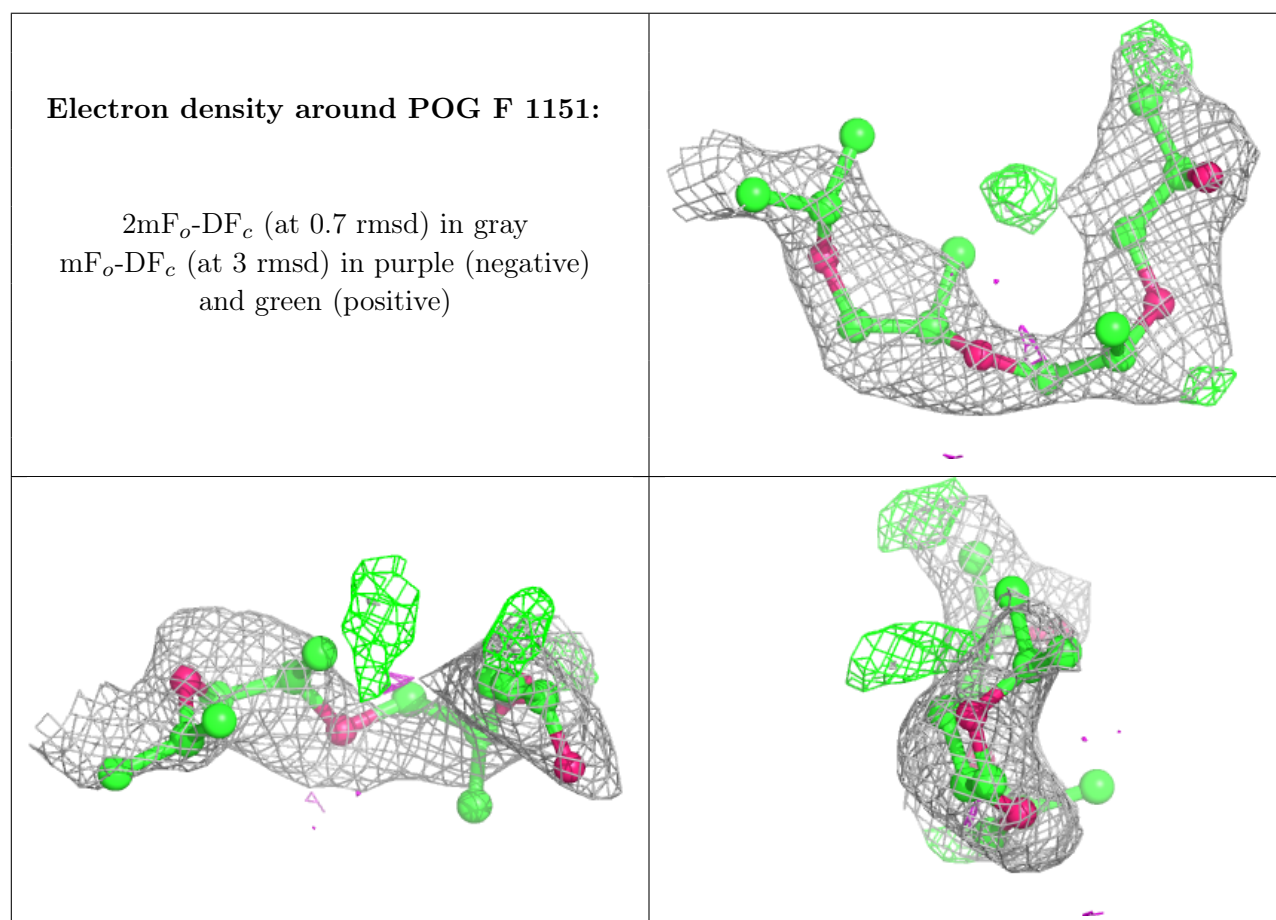
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	POG	F	1151	16/29	0.72	0.27	69,75,83,86	0
5	POG	C	1149	29/29	0.85	0.17	46,61,91,94	0
4	CL	B	1150	1/1	0.91	0.39	98,98,98,98	0
5	POG	D	1151	16/29	0.92	0.29	40,58,67,69	0
3	SO4	A	1150	5/5	0.93	0.13	46,70,76,80	0
3	SO4	F	1150	5/5	0.94	0.14	71,86,94,96	0
3	SO4	A	1149	5/5	0.96	0.15	72,72,80,86	0
3	SO4	F	1149	5/5	0.97	0.15	46,47,52,71	0
2	CMP	D	1148	22/22	0.97	0.11	27,32,33,34	0
2	CMP	E	1148	22/22	0.98	0.09	32,38,43,46	0
4	CL	A	1151	1/1	0.98	0.22	30,30,30,30	0

*Continued on next page...*

Continued from previous page...

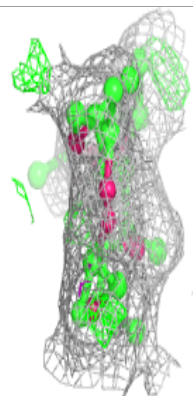
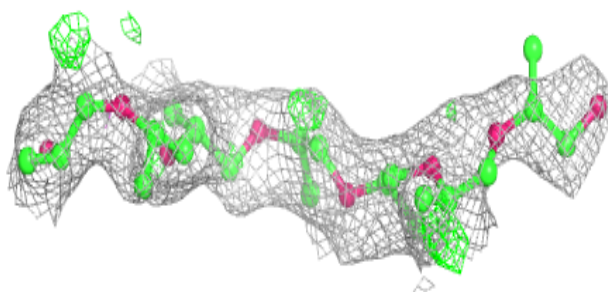
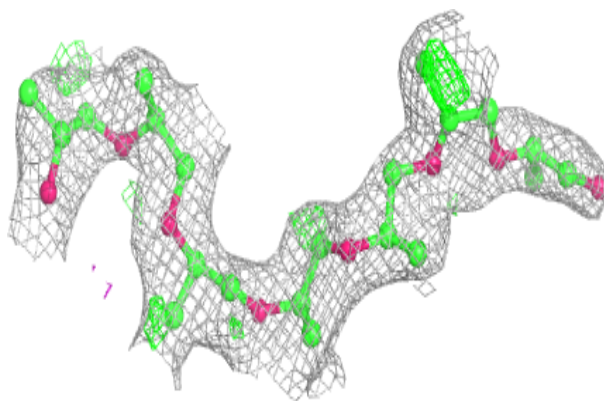
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CMP	F	1148	22/22	0.98	0.09	43,48,56,56	0
2	CMP	B	1148	22/22	0.98	0.12	27,30,32,33	0
2	CMP	C	1148	22/22	0.98	0.10	34,38,41,44	0
2	CMP	A	1148	22/22	0.98	0.11	25,35,40,44	0
3	SO4	D	1149	5/5	0.99	0.12	37,37,42,45	0
3	SO4	D	1150	5/5	0.99	0.16	37,42,47,53	0
3	SO4	B	1149	5/5	0.99	0.16	22,22,26,29	5

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.

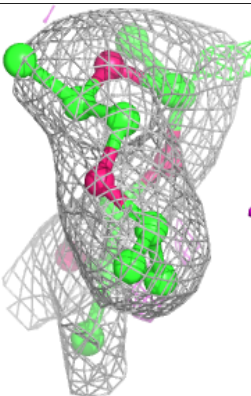
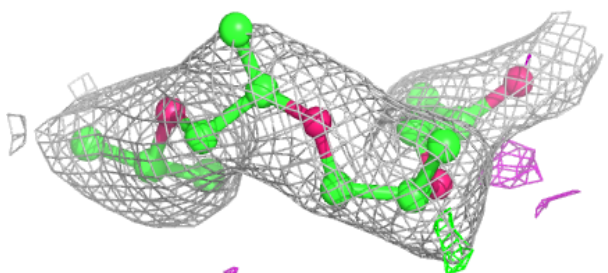
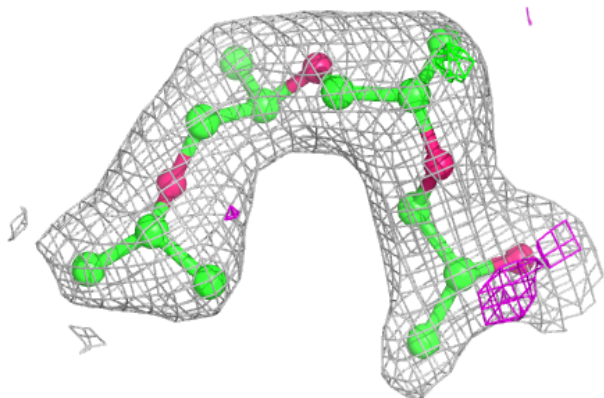


**Electron density around POG C 1149:**

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

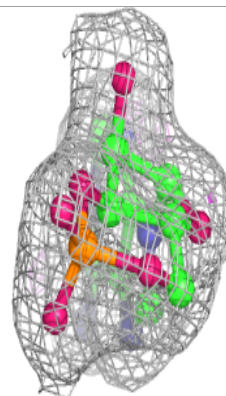
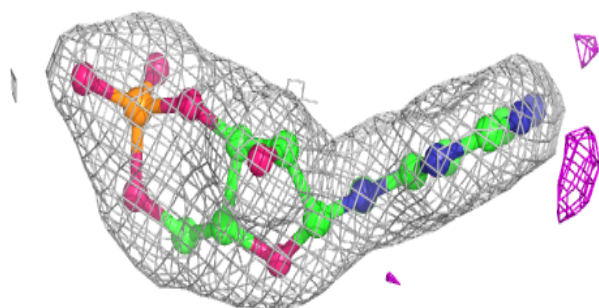
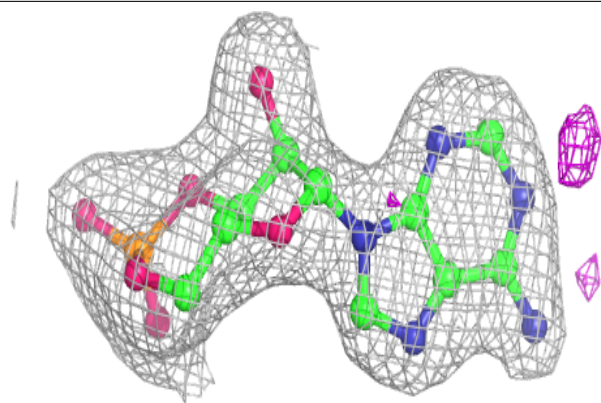
**Electron density around POG D 1151:**

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

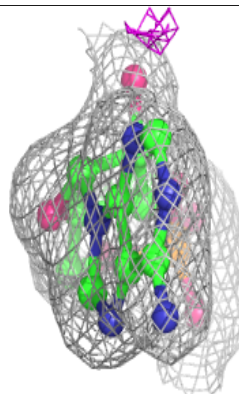
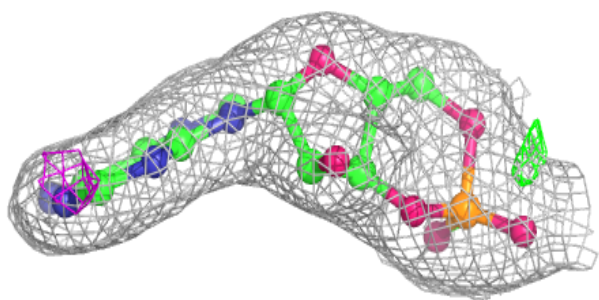
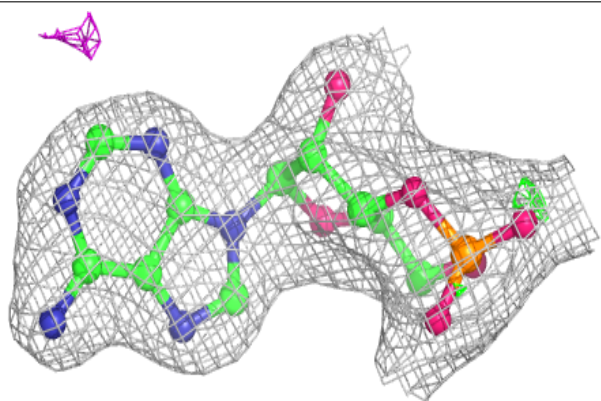


**Electron density around CMP D 1148:**

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

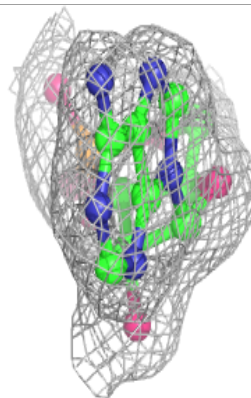
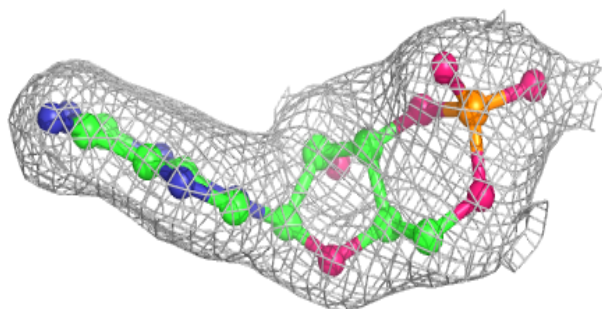
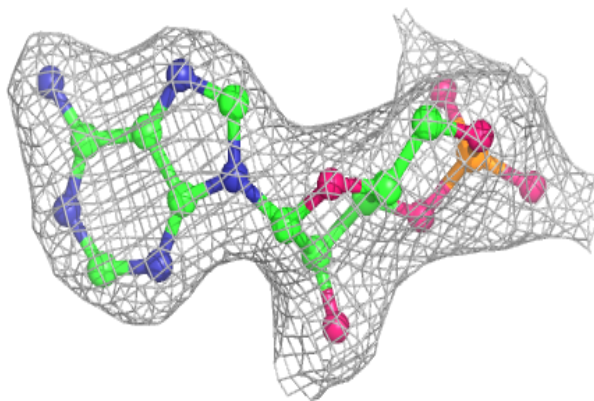
**Electron density around CMP E 1148:**

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

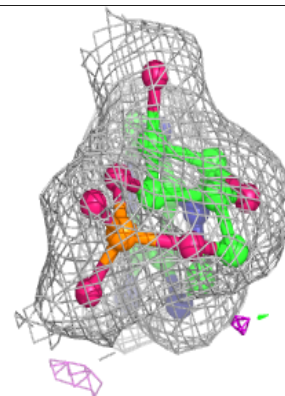
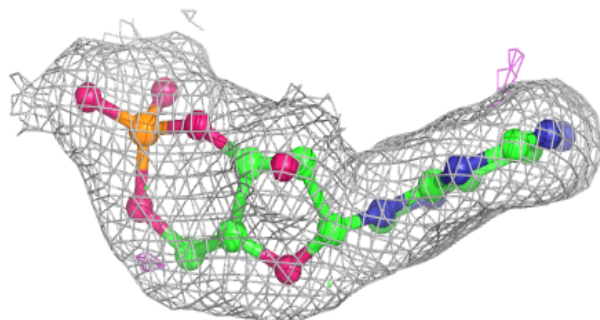
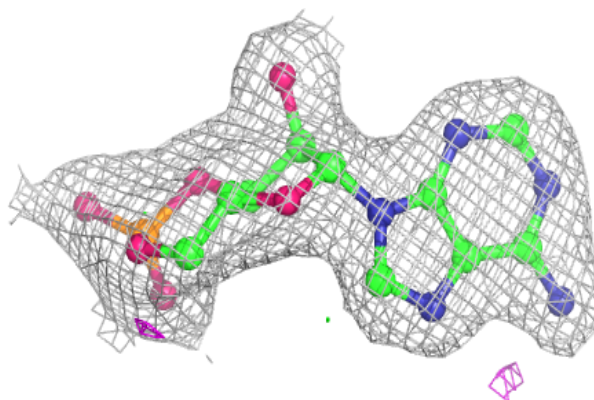


**Electron density around CMP F 1148:**

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

**Electron density around CMP B 1148:**

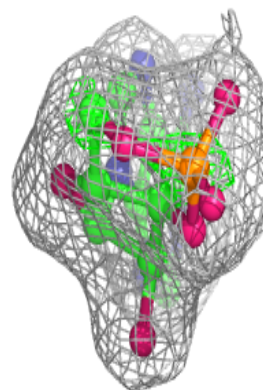
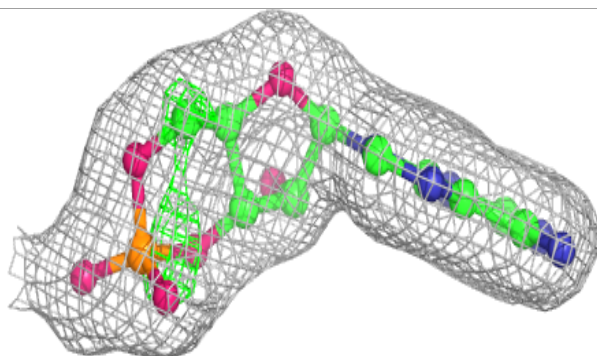
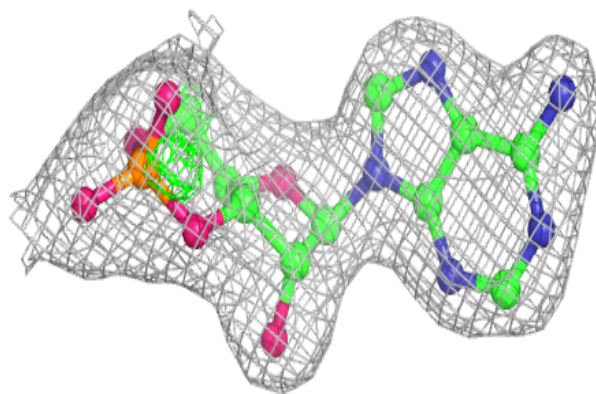
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



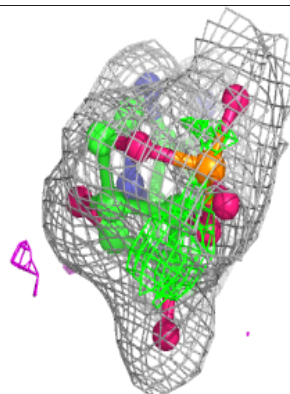
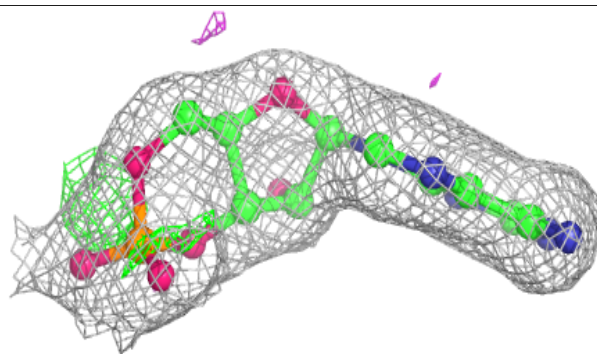
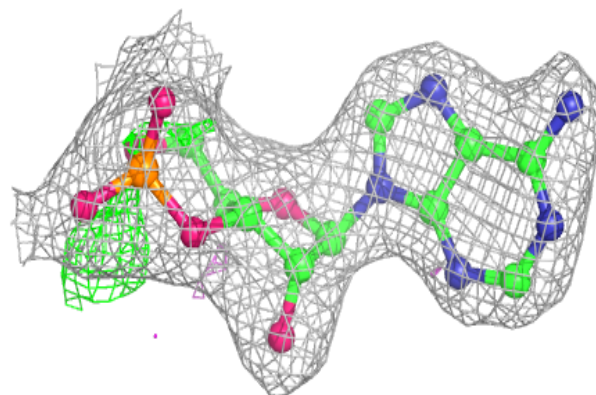


**Electron density around CMP C 1148:**

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

**Electron density around CMP A 1148:**

$2mF_o-DF_c$  (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.