



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2025 – 10:42 AM EDT

PDB ID : 9C1D / pdb\_00009c1d  
Title : Mycobacterium tuberculosis Pks13 acyltransferase incubated with DMSO  
Authors : Tang, S.; Sacchettini, J.C.; TB Structural Genomics Consortium (TBSGC)  
Deposited on : 2024-05-29  
Resolution : 2.05 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 : 4-5-2 with Phenix2.0rc1  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0rc1  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.006 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.43.1

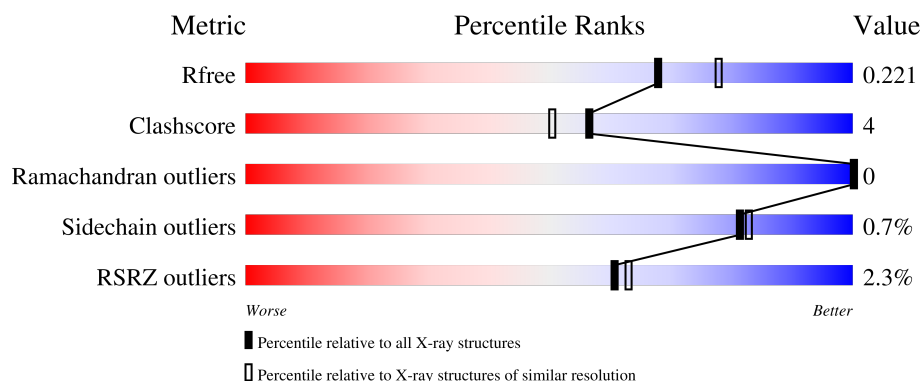
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

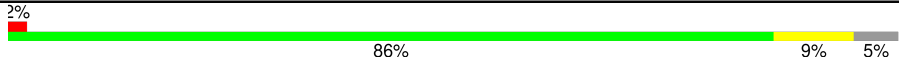
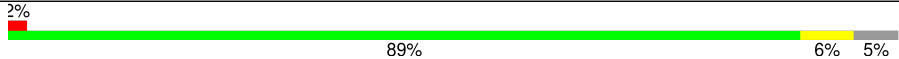
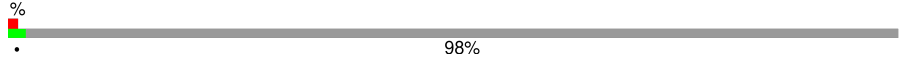
The reported resolution of this entry is 2.05 Å.

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))
$R_{free}$	164625	2096 (2.04-2.04)
Clashscore	180529	2229 (2.04-2.04)
Ramachandran outliers	177936	2217 (2.04-2.04)
Sidechain outliers	177891	2217 (2.04-2.04)
RSRZ outliers	164620	2096 (2.04-2.04)

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	491	
1	B	491	
1	C	491	

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	DMS	B	1107	-	-	X	-
6	PEG	A	1123[B]	-	-	X	-

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 8056 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 Polyketide synthase Pks13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	468	Total	C	N	O	S	0	1	0
			3538	2239	609	678	12			
1	B	468	Total	C	N	O	S	0	6	0
			3606	2274	636	684	12			
1	C	12	Total	C	N	O		0	0	0
			93	60	14	19				

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	573	SER	-	expression tag	UNP I6X8D2
A	574	ASN	-	expression tag	UNP I6X8D2
A	575	ALA	-	expression tag	UNP I6X8D2
B	573	SER	-	expression tag	UNP I6X8D2
B	574	ASN	-	expression tag	UNP I6X8D2
B	575	ALA	-	expression tag	UNP I6X8D2
C	573	SER	-	expression tag	UNP I6X8D2
C	574	ASN	-	expression tag	UNP I6X8D2
C	575	ALA	-	expression tag	UNP I6X8D2

- Molecule 2 is DIMETHYL SULFOXIDE (CCD ID: DMS) (formula: C<sub>2</sub>H<sub>6</sub>OS).



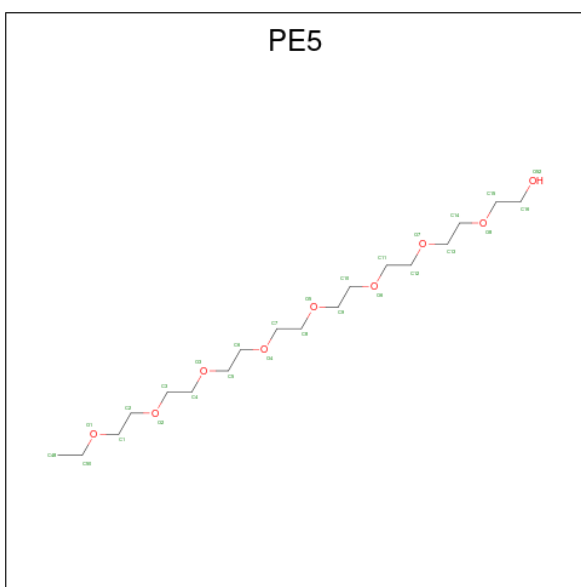
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	A	1	Total	C	O	S	0	0
			4	2	1	1		
2	B	1	Total	C	O	S	0	0
			4	2	1	1		
2	B	1	Total	C	O	S	0	0
			4	2	1	1		
2	B	1	Total	C	O	S	0	0
			4	2	1	1		
2	B	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 3 is SULFATE ION (CCD ID: SO4) (formula: O<sub>4</sub>S).



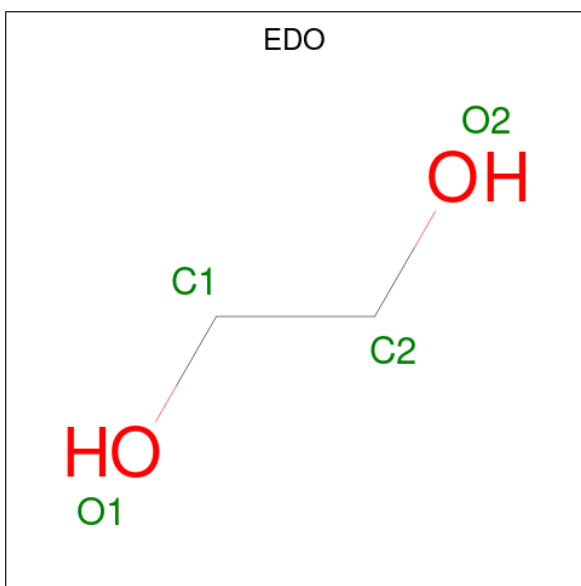
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total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

- Molecule 4 is 3,6,9,12,15,18,21,24-OCTAOXAHEXACOSAN-1-OL (CCD ID: PE5) (formula: C<sub>18</sub>H<sub>38</sub>O<sub>9</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			27	18	9		

- Molecule 5 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula:  $\text{C}_2\text{H}_6\text{O}_2$ ).



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

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 6 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			7	4	3		
6	A	1	Total	C	O	0	0
			7	4	3		
6	A	1	Total	C	O	0	0
			7	4	3		
6	A	1	Total	C	O	0	0
			7	4	3		
6	A	1	Total	C	O	0	1
			14	8	6		
6	A	1	Total	C	O	0	0
			7	4	3		
6	A	1	Total	C	O	0	0
			7	4	3		
6	B	1	Total	C	O	0	0
			7	4	3		
6	B	1	Total	C	O	0	0
			7	4	3		
6	B	1	Total	C	O	0	0
			7	4	3		

- Molecule 7 is GLYCEROL (CCD ID: GOL) (formula:  $C_3H_8O_3$ ).



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

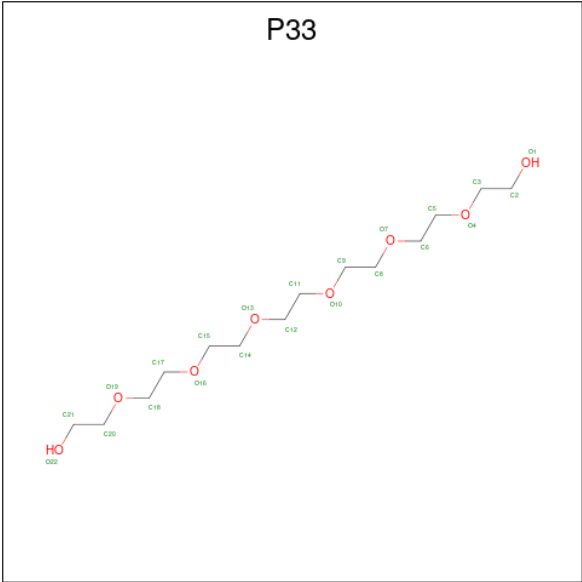
- Molecule 8 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	3	Total	Cl	0	0
			3	3		
8	B	3	Total	Cl	0	0
			3	3		

- Molecule 9 is SODIUM ION (CCD ID: NA) (formula: Na).

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

- Molecule 10 is 3,6,9,12,15,18-HEXAIOXAICOSANE-1,20-DIOL (CCD ID: P33) (formula: C<sub>14</sub>H<sub>30</sub>O<sub>8</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	B	1	Total	C O	0	0
			22	14 8		

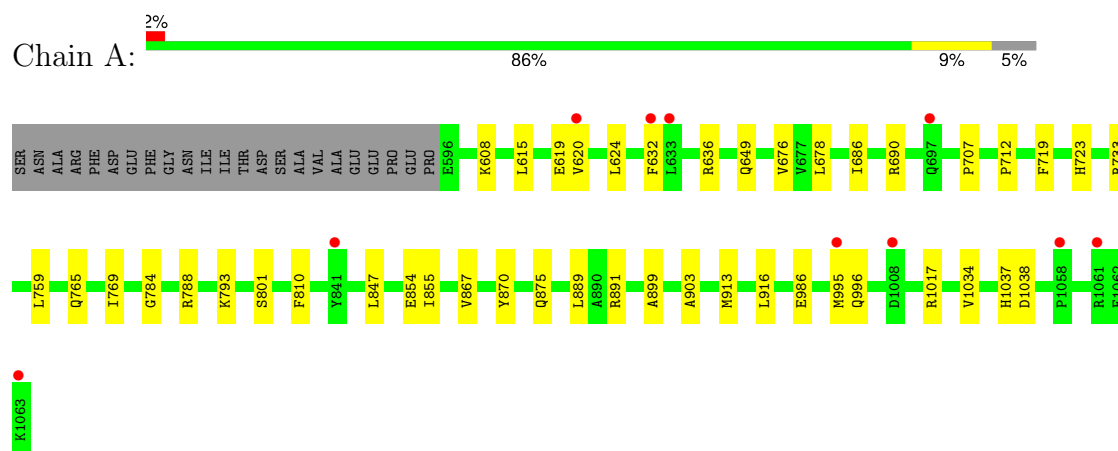
- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	270	Total	O	0	0
			270	270		
11	B	250	Total	O	0	0
			250	250		
11	C	1	Total	O	0	0
			1	1		

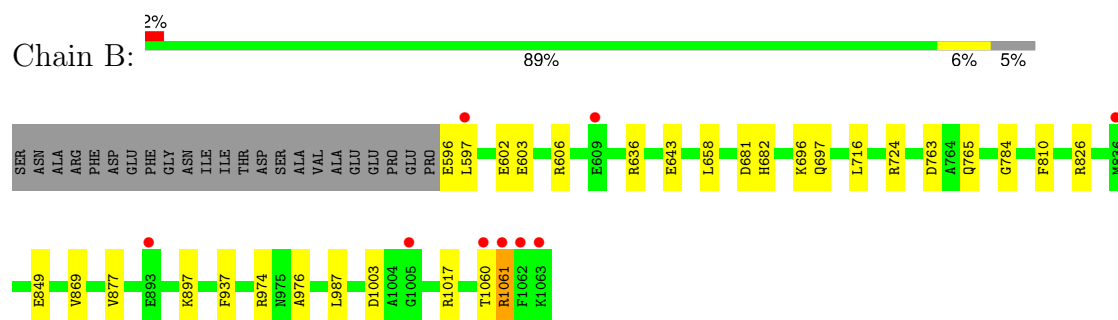
### 3 Residue-property plots [i](#)

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.

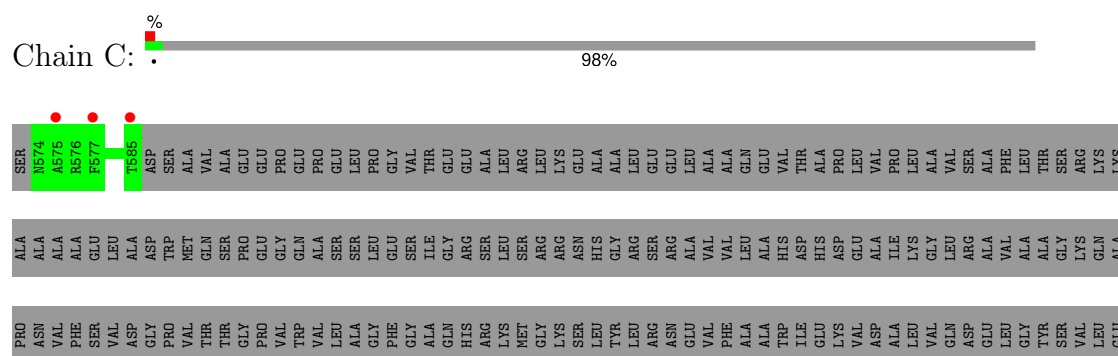
#### • Molecule 1: Polyketide synthase Pks13



#### • Molecule 1: Polyketide synthase Pks13



#### • Molecule 1: Polyketide synthase Pks13



[illegible]

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.18Å 106.18Å 259.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.14 – 2.05 49.14 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.3 (49.14-2.05) 99.3 (49.14-2.05)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.13 (at 2.05Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, $R_{free}$	0.191 , 0.221 0.191 , 0.221	Depositor DCC
$R_{free}$ test set	91098 reflections (2.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.0	Xtriage
Anisotropy	0.494	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 38.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8056	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.97% 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: PEG, CL, PE5, P33, NA, EDO, GOL, DMS, SO4

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	0.15	0/3612	0.32	0/4910
1	B	0.15	0/3680	0.32	0/4996
1	C	0.11	0/94	0.27	0/126
All	All	0.15	0/7386	0.32	0/10032

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [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	3538	0	3482	41	0
1	B	3606	0	3565	24	0
1	C	93	0	80	0	0
2	A	12	0	18	1	0
2	B	20	0	30	6	0
3	A	30	0	0	1	0
3	B	20	0	0	1	0
3	C	5	0	0	0	0
4	A	27	0	38	6	0
5	A	40	0	60	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	32	0	48	1	0
6	A	56	0	80	13	0
6	B	21	0	30	2	0
7	A	6	0	8	2	0
8	A	3	0	0	0	0
8	B	3	0	0	1	0
9	A	1	0	0	0	0
10	B	22	0	30	0	0
11	A	270	0	0	5	0
11	B	250	0	0	2	0
11	C	1	0	0	0	0
All	All	8056	0	7469	67	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 4.

All (67) 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:1038:ASP:HB2	6:A:1123[B]:PEG:H42	1.52	0.90
1:A:793:LYS:H	6:A:1123[B]:PEG:H41	1.42	0.84
1:A:889:LEU:HD22	1:A:899:ALA:HB1	1.71	0.73
1:A:996:GLN:NE2	2:A:1101:DMS:S	2.59	0.72
1:B:1017:ARG:HD2	2:B:1110:DMS:H12	1.72	0.72
1:A:1037:HIS:HA	6:A:1123[A]:PEG:H22	1.72	0.71
1:A:712:PRO:HD3	6:A:1123[B]:PEG:H21	1.72	0.70
1:A:707:PRO:HB3	6:A:1122:PEG:H21	1.75	0.68
1:B:602:GLU:H	5:B:1121:EDO:H12	1.57	0.68
1:A:875:GLN:OE1	11:A:1201:HOH:O	2.12	0.67
1:A:632:PHE:HB3	1:A:636:ARG:HH12	1.60	0.66
1:A:769:ILE:HG23	4:A:1108:PE5:H61	1.80	0.64
1:A:1038:ASP:H	6:A:1123[A]:PEG:H31	1.62	0.64
2:B:1110:DMS:H13	8:B:1122:CL:CL	2.35	0.64
1:A:1037:HIS:HA	6:A:1123[B]:PEG:H12	1.81	0.63
1:A:788:ARG:HD2	7:A:1124:GOL:H11	1.82	0.61
1:A:615:LEU:HD13	5:A:1111:EDO:H11	1.84	0.60
1:A:649:GLN:HG2	1:A:686:ILE:HD13	1.86	0.57
1:A:636:ARG:NH2	11:A:1206:HOH:O	2.38	0.57
1:A:854:GLU:HA	6:A:1125:PEG:H12	1.86	0.56
1:B:603:GLU:OE2	1:B:606:ARG:NH2	2.36	0.54
1:B:681:ASP:HB2	6:B:1120:PEG:H22	1.90	0.54

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:765:GLN:NE2	11:B:1206:HOH:O	2.30	0.53
1:A:916:LEU:HD21	4:A:1108:PE5:H151	1.92	0.52
1:B:682:HIS:HB2	6:B:1120:PEG:H11	1.92	0.51
1:A:707:PRO:HB3	6:A:1122:PEG:H41	1.92	0.50
1:A:733:ARG:NH2	11:A:1210:HOH:O	2.44	0.50
1:A:1017:ARG:HD2	5:A:1109:EDO:H12	1.94	0.50
1:B:1060:THR:OG1	1:B:1061[B]:ARG:NH1	2.44	0.49
1:B:869:VAL:HB	1:B:877:VAL:HB	1.94	0.49
1:A:810:PHE:HD1	7:A:1124:GOL:H12	1.78	0.49
1:A:608:LYS:NZ	5:A:1114:EDO:O1	2.45	0.48
1:B:849:GLU:HB3	1:B:897:LYS:HB3	1.95	0.48
1:B:696:LYS:HE3	2:B:1107:DMS:H12	1.95	0.48
1:A:723:HIS:HE1	5:A:1110:EDO:H22	1.78	0.48
1:A:784:GLY:HA3	1:A:810:PHE:CZ	2.49	0.47
1:B:697:GLN:HG3	2:B:1107:DMS:H21	1.96	0.47
1:A:913:MET:HE1	4:A:1108:PE5:H112	1.97	0.47
1:A:903:ALA:H	4:A:1108:PE5:H483	1.80	0.46
1:A:719:PHE:HE1	4:A:1108:PE5:H71	1.80	0.45
1:B:826:ARG:NH2	3:B:1101:SO4:O1	2.39	0.45
1:A:707:PRO:CB	6:A:1122:PEG:H21	2.45	0.45
1:A:759:LEU:HD22	1:A:765:GLN:HG2	1.98	0.45
1:A:847:LEU:HD11	1:A:875:GLN:HB3	1.99	0.45
1:B:696:LYS:HA	2:B:1107:DMS:H23	1.99	0.45
1:A:624:LEU:HD11	1:A:678:LEU:HB3	1.98	0.44
1:B:596:GLU:HG3	1:B:597:LEU:HD13	1.99	0.44
1:A:891:ARG:NH1	11:A:1208:HOH:O	2.42	0.44
1:B:636:ARG:HB3	1:B:1061[B]:ARG:NH2	2.33	0.44
1:B:724:ARG:HD3	1:B:763:ASP:O	2.18	0.43
1:B:636:ARG:HD2	11:B:1435:HOH:O	2.18	0.43
1:B:697:GLN:H	2:B:1107:DMS:H11	1.83	0.43
1:A:676:VAL:HG23	1:A:1034[B]:VAL:HG13	2.01	0.43
1:A:1038:ASP:CB	6:A:1123[B]:PEG:H42	2.36	0.43
1:A:801:SER:HB3	3:A:1102:SO4:O2	2.18	0.42
1:B:716:LEU:HD11	1:B:987:LEU:HD12	2.01	0.42
4:A:1108:PE5:H151	4:A:1108:PE5:H132	1.78	0.42
1:A:867:VAL:HG11	1:A:870:TYR:CZ	2.55	0.42
1:A:686:ILE:O	1:A:690:ARG:HG2	2.20	0.42
1:B:974[B]:ARG:NH2	1:B:1003:ASP:OD2	2.45	0.41
1:B:784:GLY:HA3	1:B:810:PHE:CZ	2.56	0.41
1:A:854:GLU:HG2	6:A:1125:PEG:H21	2.03	0.41
1:B:643:GLU:OE2	1:B:1060:THR:HA	2.21	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:937:PHE:CZ	1:B:976:ALA:HA	2.56	0.41
1:B:658:LEU:HD23	1:B:658:LEU:HA	1.96	0.40
1:A:793:LYS:N	6:A:1123[B]:PEG:H41	2.22	0.40
1:A:986:GLU:OE2	11:A:1202:HOH:O	2.22	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	467/491 (95%)	461 (99%)	6 (1%)	0	100	100
1	B	472/491 (96%)	465 (98%)	7 (2%)	0	100	100
1	C	10/491 (2%)	10 (100%)	0	0	100	100
All	All	949/1473 (64%)	936 (99%)	13 (1%)	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	362/387 (94%)	358 (99%)	4 (1%)	70	71
1	B	370/387 (96%)	368 (100%)	2 (0%)	86	88
1	C	9/387 (2%)	9 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	741/1161 (64%)	735 (99%)	6 (1%)	81	80

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

Mol	Chain	Res	Type
1	A	619	GLU
1	A	620	VAL
1	A	855	ILE
1	A	995	MET
1	B	1061[A]	ARG
1	B	1061[B]	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	618	GLN
1	A	649	GLN
1	A	780	GLN
1	A	1031	GLN
1	B	697	GLN
1	B	780	GLN
1	B	789	HIS
1	B	1031	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry

Of 58 ligands modelled in this entry, 7 are monoatomic - leaving 51 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
5	EDO	B	1119	-	3,3,3	0.42	0	2,2,2	0.38	0
3	SO4	A	1105	-	4,4,4	0.24	0	6,6,6	0.07	0
5	EDO	B	1121	-	3,3,3	0.44	0	2,2,2	0.39	0
6	PEG	B	1111	-	6,6,6	0.13	0	5,5,5	0.08	0
4	PE5	A	1108	-	26,26,26	0.34	0	25,25,25	0.34	0
5	EDO	A	1110	-	3,3,3	0.41	0	2,2,2	0.42	0
2	DMS	A	1116	-	3,3,3	0.65	0	3,3,3	0.56	0
6	PEG	B	1120	-	6,6,6	0.16	0	5,5,5	0.05	0
10	P33	B	1105	-	21,21,21	0.30	0	20,20,20	0.17	0
5	EDO	A	1126	-	3,3,3	0.41	0	2,2,2	0.36	0
6	PEG	A	1117	-	6,6,6	0.12	0	5,5,5	0.10	0
5	EDO	A	1114	-	3,3,3	0.44	0	2,2,2	0.40	0
2	DMS	A	1118	-	3,3,3	0.66	0	3,3,3	0.53	0
3	SO4	A	1107	-	4,4,4	0.23	0	6,6,6	0.08	0
7	GOL	A	1124	-	5,5,5	0.93	0	5,5,5	1.14	1 (20%)
3	SO4	B	1102	-	4,4,4	0.23	0	6,6,6	0.08	0
5	EDO	A	1128	-	3,3,3	0.43	0	2,2,2	0.39	0
5	EDO	B	1116	-	3,3,3	0.43	0	2,2,2	0.38	0
2	DMS	B	1107	-	3,3,3	0.66	0	3,3,3	0.62	0
5	EDO	A	1115	-	3,3,3	0.51	0	2,2,2	0.26	0
5	EDO	B	1117	-	3,3,3	0.43	0	2,2,2	0.44	0
3	SO4	B	1104	-	4,4,4	0.24	0	6,6,6	0.06	0
2	DMS	B	1114	-	3,3,3	0.66	0	3,3,3	0.52	0
2	DMS	B	1109	-	3,3,3	0.66	0	3,3,3	0.54	0
2	DMS	A	1101	-	3,3,3	0.68	0	3,3,3	0.53	0
3	SO4	B	1103	-	4,4,4	0.21	0	6,6,6	0.05	0
6	PEG	A	1122	-	6,6,6	0.10	0	5,5,5	0.13	0
5	EDO	A	1120	-	3,3,3	0.44	0	2,2,2	0.33	0
5	EDO	B	1112	-	3,3,3	0.44	0	2,2,2	0.37	0
6	PEG	A	1121	-	6,6,6	0.12	0	5,5,5	0.09	0
5	EDO	A	1112	-	3,3,3	0.45	0	2,2,2	0.37	0
5	EDO	A	1111	-	3,3,3	0.47	0	2,2,2	0.32	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	SO4	A	1106	-	4,4,4	0.23	0	6,6,6	0.11	0
6	PEG	A	1123[A]	-	6,6,6	0.12	0	5,5,5	0.05	0
5	EDO	B	1108	-	3,3,3	0.41	0	2,2,2	0.49	0
6	PEG	A	1123[B]	-	6,6,6	0.14	0	5,5,5	0.24	0
5	EDO	B	1118	-	3,3,3	0.41	0	2,2,2	0.50	0
5	EDO	A	1119	-	3,3,3	0.46	0	2,2,2	0.39	0
3	SO4	A	1104	-	4,4,4	0.24	0	6,6,6	0.11	0
6	PEG	A	1125	-	6,6,6	0.11	0	5,5,5	0.09	0
2	DMS	B	1110	-	3,3,3	0.66	0	3,3,3	0.62	0
3	SO4	A	1103	-	4,4,4	0.24	0	6,6,6	0.09	0
6	PEG	A	1127	-	6,6,6	0.13	0	5,5,5	0.10	0
5	EDO	A	1109	-	3,3,3	0.44	0	2,2,2	0.33	0
5	EDO	B	1106	-	3,3,3	0.44	0	2,2,2	0.35	0
6	PEG	B	1115	-	6,6,6	0.10	0	5,5,5	0.13	0
6	PEG	A	1113	-	6,6,6	0.10	0	5,5,5	0.07	0
2	DMS	B	1113	-	3,3,3	0.68	0	3,3,3	0.55	0
3	SO4	B	1101	-	4,4,4	0.25	0	6,6,6	0.27	0
3	SO4	A	1102	-	4,4,4	0.23	0	6,6,6	0.12	0
3	SO4	C	1101	-	4,4,4	0.24	0	6,6,6	0.07	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
5	EDO	B	1119	-	-	0/1/1/1	-
5	EDO	B	1121	-	-	0/1/1/1	-
6	PEG	B	1111	-	-	0/4/4/4	-
4	PE5	A	1108	-	-	13/24/24/24	-
5	EDO	A	1110	-	-	0/1/1/1	-
10	P33	B	1105	-	-	6/19/19/19	-
6	PEG	B	1120	-	-	2/4/4/4	-
5	EDO	A	1126	-	-	0/1/1/1	-
6	PEG	A	1117	-	-	2/4/4/4	-
5	EDO	A	1114	-	-	0/1/1/1	-
7	GOL	A	1124	-	-	0/4/4/4	-
5	EDO	A	1128	-	-	0/1/1/1	-
5	EDO	B	1116	-	-	0/1/1/1	-
5	EDO	A	1115	-	-	0/1/1/1	-
5	EDO	B	1117	-	-	1/1/1/1	-
6	PEG	A	1122	-	-	1/4/4/4	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	EDO	A	1120	-	-	0/1/1/1	-
5	EDO	B	1112	-	-	0/1/1/1	-
6	PEG	A	1121	-	-	1/4/4/4	-
5	EDO	A	1111	-	-	0/1/1/1	-
6	PEG	A	1123[A]	-	-	3/4/4/4	-
5	EDO	B	1108	-	-	0/1/1/1	-
6	PEG	A	1123[B]	-	-	1/4/4/4	-
5	EDO	B	1118	-	-	0/1/1/1	-
5	EDO	A	1119	-	-	0/1/1/1	-
6	PEG	A	1125	-	-	2/4/4/4	-
6	PEG	A	1127	-	-	1/4/4/4	-
5	EDO	A	1109	-	-	1/1/1/1	-
5	EDO	B	1106	-	-	1/1/1/1	-
6	PEG	B	1115	-	-	1/4/4/4	-
6	PEG	A	1113	-	-	2/4/4/4	-
5	EDO	A	1112	-	-	1/1/1/1	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	1124	GOL	C3-C2-C1	-2.03	104.34	111.80

There are no chirality outliers.

All (39) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1108	PE5	O3-C5-C6-O4
4	A	1108	PE5	O6-C10-C9-O5
4	A	1108	PE5	O7-C13-C14-O8
6	A	1113	PEG	O2-C3-C4-O4
6	A	1123[B]	PEG	O1-C1-C2-O2
6	B	1120	PEG	O1-C1-C2-O2
6	A	1125	PEG	O2-C3-C4-O4
6	A	1113	PEG	C4-C3-O2-C2
5	A	1109	EDO	O1-C1-C2-O2
4	A	1108	PE5	C13-C14-O8-C15
5	B	1117	EDO	O1-C1-C2-O2
6	A	1127	PEG	O2-C3-C4-O4
4	A	1108	PE5	C48-C50-O1-C1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
6	A	1121	PEG	O2-C3-C4-O4
6	A	1117	PEG	O1-C1-C2-O2
6	A	1122	PEG	O1-C1-C2-O2
4	A	1108	PE5	C9-C10-O6-C11
10	B	1105	P33	C21-C20-O19-C18
4	A	1108	PE5	C10-C9-O5-C8
4	A	1108	PE5	C6-C5-O3-C4
6	A	1117	PEG	O2-C3-C4-O4
6	B	1115	PEG	C1-C2-O2-C3
4	A	1108	PE5	O2-C3-C4-O3
4	A	1108	PE5	C1-C2-O2-C3
5	A	1112	EDO	O1-C1-C2-O2
4	A	1108	PE5	C5-C6-O4-C7
6	A	1123[A]	PEG	C4-C3-O2-C2
10	B	1105	P33	C17-C18-O19-C20
6	A	1123[A]	PEG	O1-C1-C2-O2
5	B	1106	EDO	O1-C1-C2-O2
6	A	1123[A]	PEG	C1-C2-O2-C3
10	B	1105	P33	C8-C9-O10-C11
10	B	1105	P33	C6-C5-O4-C3
4	A	1108	PE5	O8-C15-C16-O52
6	B	1120	PEG	O2-C3-C4-O4
6	A	1125	PEG	O1-C1-C2-O2
10	B	1105	P33	O13-C14-C15-O16
10	B	1105	P33	O16-C17-C18-O19
4	A	1108	PE5	O6-C11-C12-O7

There are no ring outliers.

17 monomers are involved in 37 short contacts:

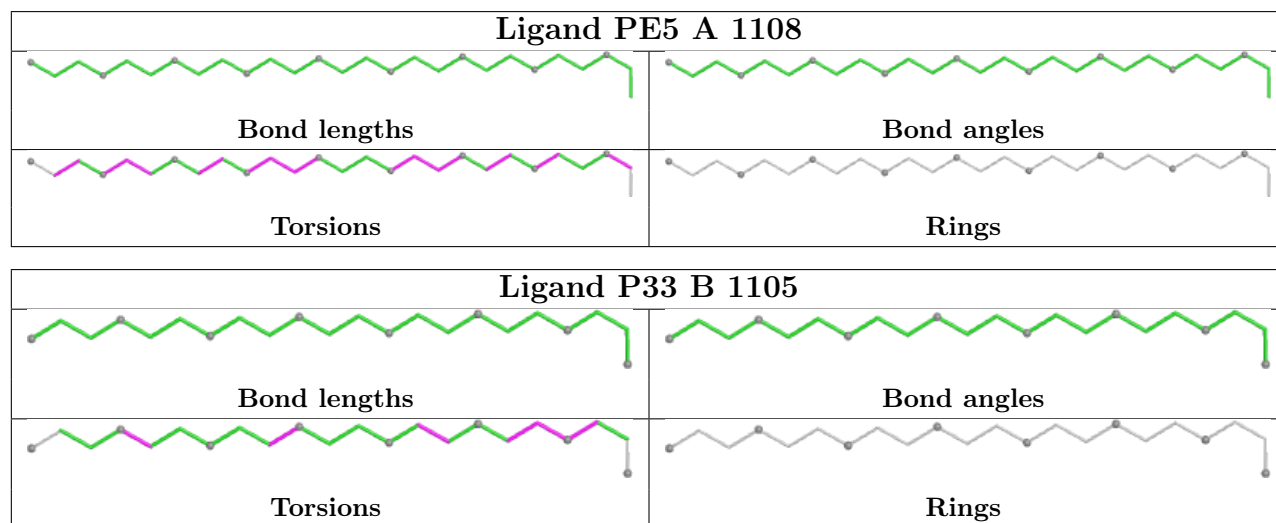
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1121	EDO	1	0
4	A	1108	PE5	6	0
5	A	1110	EDO	1	0
6	B	1120	PEG	2	0
5	A	1114	EDO	1	0
7	A	1124	GOL	2	0
2	B	1107	DMS	4	0
2	A	1101	DMS	1	0
6	A	1122	PEG	3	0
5	A	1111	EDO	1	0
6	A	1123[A]	PEG	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1123[B]	PEG	6	0
6	A	1125	PEG	2	0
2	B	1110	DMS	2	0
5	A	1109	EDO	1	0
3	B	1101	SO4	1	0
3	A	1102	SO4	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	468/491 (95%)	-0.03	10 (2%) 63 66	18, 36, 52, 79	1 (0%)
1	B	468/491 (95%)	0.14	9 (1%) 66 69	15, 37, 55, 80	6 (1%)
1	C	12/491 (2%)	1.22	3 (25%) 2 1	42, 55, 82, 87	0
All	All	948/1473 (64%)	0.07	22 (2%) 61 63	15, 37, 55, 87	7 (0%)

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1061[A]	ARG	5.4
1	B	1062	PHE	4.1
1	C	585	THR	3.2
1	B	1063	LYS	3.2
1	A	841	TYR	3.0
1	A	620	VAL	2.8
1	C	577	PHE	2.8
1	B	1060	THR	2.8
1	B	836	MET	2.7
1	A	633	LEU	2.3
1	A	697	GLN	2.3
1	A	1061	ARG	2.2
1	A	1063	LYS	2.2
1	B	1005	GLY	2.2
1	B	893	GLU	2.1
1	A	1058	PRO	2.1
1	A	1008	ASP	2.1
1	B	597	LEU	2.1
1	C	575	ALA	2.1
1	B	609	GLU	2.1
1	A	632	PHE	2.1
1	A	995	MET	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( $\text{\AA}^2$ )	Q<0.9
6	PEG	A	1123[A]	7/7	0.57	0.29	34,40,43,43	7
6	PEG	A	1123[B]	7/7	0.57	0.29	28,35,44,45	7
2	DMS	B	1114	4/4	0.69	0.30	65,72,77,79	0
2	DMS	A	1116	4/4	0.71	0.23	58,66,71,83	0
2	DMS	B	1107	4/4	0.72	0.19	58,59,80,87	0
5	EDO	B	1121	4/4	0.72	0.22	51,55,56,58	0
7	GOL	A	1124	6/6	0.74	0.17	43,50,56,63	0
5	EDO	A	1119	4/4	0.75	0.28	53,53,53,62	0
5	EDO	A	1120	4/4	0.75	0.18	48,55,60,61	0
5	EDO	A	1110	4/4	0.76	0.22	50,52,53,54	0
2	DMS	B	1109	4/4	0.76	0.19	76,77,91,97	0
6	PEG	A	1122	7/7	0.76	0.17	41,50,53,64	0
5	EDO	A	1114	4/4	0.77	0.17	55,59,61,63	0
2	DMS	A	1118	4/4	0.77	0.22	74,75,86,90	0
3	SO4	B	1102	5/5	0.78	0.09	84,89,92,94	0
5	EDO	B	1117	4/4	0.79	0.16	52,57,57,61	0
5	EDO	A	1115	4/4	0.79	0.20	42,45,46,60	0
5	EDO	B	1116	4/4	0.79	0.21	46,47,51,51	0
5	EDO	A	1111	4/4	0.80	0.21	53,56,59,59	0
6	PEG	B	1115	7/7	0.80	0.17	51,58,62,68	0
3	SO4	A	1107	5/5	0.80	0.19	53,58,61,67	5
5	EDO	B	1108	4/4	0.81	0.17	45,58,58,64	0
6	PEG	B	1120	7/7	0.81	0.18	47,49,55,59	0
5	EDO	B	1119	4/4	0.81	0.20	53,57,58,59	0
8	CL	B	1123	1/1	0.81	0.16	78,78,78,78	0
3	SO4	A	1106	5/5	0.83	0.15	43,52,61,63	5
3	SO4	B	1103	5/5	0.83	0.13	42,52,70,74	0

*Continued on next page...*

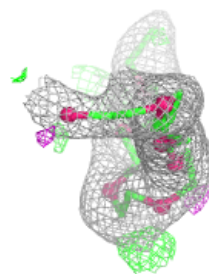
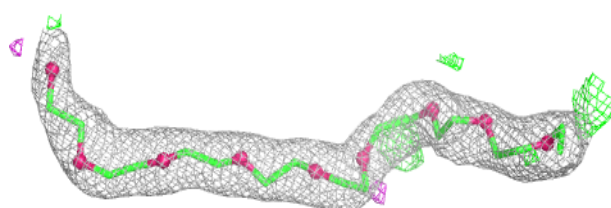
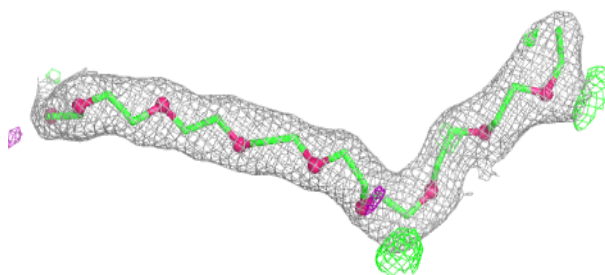
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	SO4	C	1101	5/5	0.83	0.09	65,69,75,79	0
5	EDO	B	1106	4/4	0.83	0.24	38,47,52,59	0
5	EDO	B	1112	4/4	0.84	0.15	50,56,60,60	0
6	PEG	A	1125	7/7	0.84	0.17	53,55,59,65	0
8	CL	B	1122	1/1	0.84	0.15	74,74,74,74	0
6	PEG	A	1121	7/7	0.84	0.15	47,53,60,63	0
5	EDO	A	1109	4/4	0.86	0.24	40,42,53,63	0
2	DMS	A	1101	4/4	0.86	0.21	39,44,47,63	4
5	EDO	B	1118	4/4	0.86	0.15	36,49,49,50	0
8	CL	A	1131	1/1	0.87	0.10	88,88,88,88	0
2	DMS	B	1110	4/4	0.87	0.21	49,61,65,68	0
3	SO4	A	1105	5/5	0.87	0.14	42,49,51,53	5
5	EDO	A	1128	4/4	0.88	0.16	42,48,60,64	0
2	DMS	B	1113	4/4	0.88	0.19	59,59,61,74	0
6	PEG	A	1127	7/7	0.88	0.13	44,54,61,61	0
6	PEG	B	1111	7/7	0.88	0.12	43,51,57,60	0
5	EDO	A	1126	4/4	0.88	0.14	54,55,58,66	0
6	PEG	A	1113	7/7	0.90	0.12	40,47,54,55	0
6	PEG	A	1117	7/7	0.90	0.12	51,59,61,67	0
3	SO4	A	1103	5/5	0.90	0.08	67,67,74,78	0
4	PE5	A	1108	27/27	0.91	0.13	25,43,59,64	0
3	SO4	B	1104	5/5	0.91	0.10	42,43,45,45	5
3	SO4	A	1104	5/5	0.91	0.10	48,53,61,65	0
8	CL	B	1124	1/1	0.91	0.12	75,75,75,75	0
8	CL	A	1129	1/1	0.92	0.09	74,74,74,74	0
8	CL	A	1130	1/1	0.92	0.10	64,64,64,64	0
5	EDO	A	1112	4/4	0.92	0.13	45,48,53,55	0
10	P33	B	1105	22/22	0.93	0.11	35,45,53,58	0
9	NA	A	1132	1/1	0.94	0.13	46,46,46,46	0
3	SO4	B	1101	5/5	0.96	0.11	42,42,45,45	0
3	SO4	A	1102	5/5	0.98	0.05	34,35,39,41	0

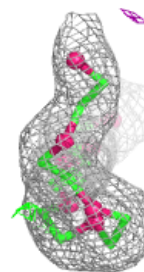
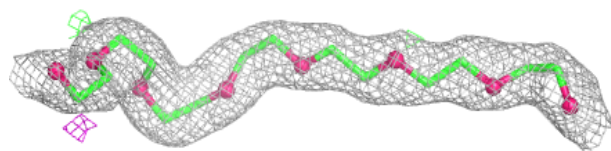
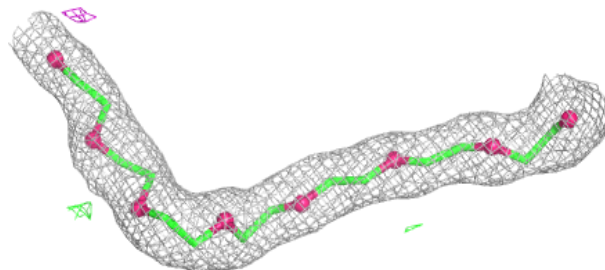
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around PE5 A 1108:**

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

**Electron density around P33 B 1105:**

$2mF_o-DF_c$  (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.