



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 03:27 pm BST

PDB ID : 6I7S  
Title : Microsomal triglyceride transfer protein  
Authors : Biterova, E.; Isupov, M.N.; Keegan, R.M.; Lebedev, A.A.; Ruddock, L.W.  
Deposited on : 2018-11-17  
Resolution : 2.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

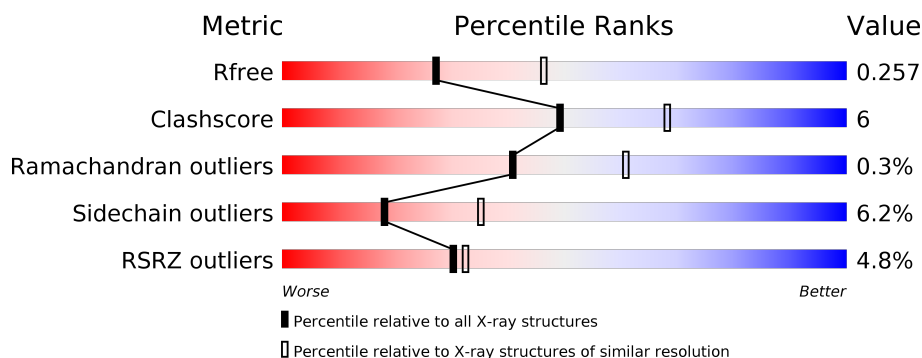
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 on 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	492	<div> <div>14%</div> <div> <div></div> <div>80%</div> <div>13%</div> <div>6%</div> </div> </div>
1	B	492	<div> <div>5%</div> <div> <div></div> <div>78%</div> <div>14%</div> <div>6%</div> </div> </div>
2	G	884	<div> <div>%</div> <div> <div></div> <div>82%</div> <div>15%</div> <div>••</div> </div> </div>
2	H	884	<div> <div>3%</div> <div> <div></div> <div>81%</div> <div>14%</div> <div>••</div> </div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	PGE	B	602	-	-	-	X
5	PGE	B	603	-	-	-	X
7	PEG	B	609	-	-	-	X
7	PEG	H	916	-	-	-	X
7	PEG	H	917	-	-	X	-
7	PEG	H	918	-	-	-	X
8	PE5	G	901	-	-	-	X
8	PE5	H	901	-	-	-	X

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 21452 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 Protein disulfide-isomerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	461	Total	C	N	O	S	0	1	0
			3668	2351	602	706	9			
1	B	462	Total	C	N	O	S	0	1	0
			3676	2354	600	713	9			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	17	MET	-	initiating methionine	UNP P07237
B	17	MET	-	initiating methionine	UNP P07237

- Molecule 2 is a protein called Microsomal triglyceride transfer protein large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	863	Total	C	N	O	S	0	4	0
			6787	4299	1164	1296	28			
2	H	862	Total	C	N	O	S	0	2	0
			6757	4277	1162	1290	28			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	11	MET	-	initiating methionine	UNP P55157
G	12	HIS	-	expression tag	UNP P55157
G	13	HIS	-	expression tag	UNP P55157
G	14	HIS	-	expression tag	UNP P55157
G	15	HIS	-	expression tag	UNP P55157
G	16	HIS	-	expression tag	UNP P55157
G	17	HIS	-	expression tag	UNP P55157
G	18	MET	-	expression tag	UNP P55157
H	11	MET	-	initiating methionine	UNP P55157
H	12	HIS	-	expression tag	UNP P55157

*Continued on next page...*

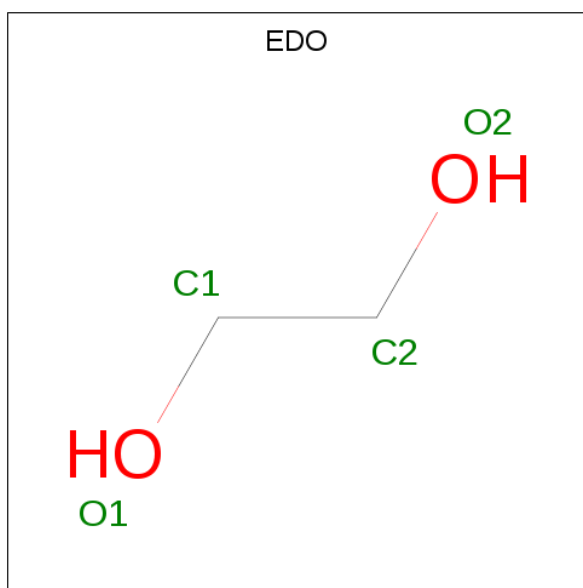
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	13	HIS	-	expression tag	UNP P55157
H	14	HIS	-	expression tag	UNP P55157
H	15	HIS	-	expression tag	UNP P55157
H	16	HIS	-	expression tag	UNP P55157
H	17	HIS	-	expression tag	UNP P55157
H	18	MET	-	expression tag	UNP P55157

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Ca 1 1	0	0
3	A	1	Total Ca 1 1	0	0

- 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
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0

*Continued on next page...*

*Continued from previous page...*

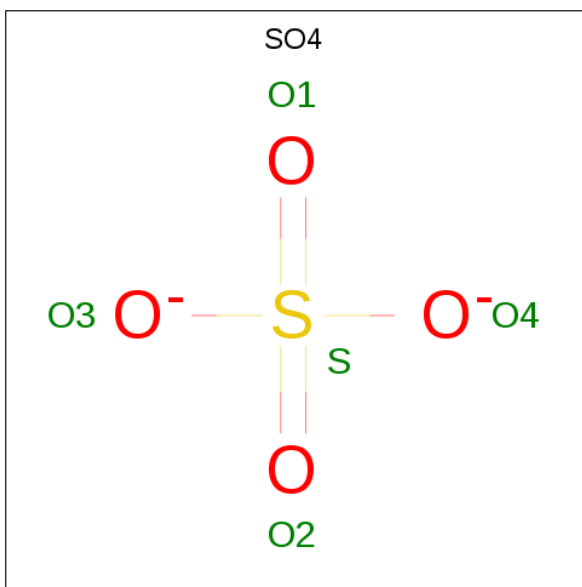
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	G	1	Total	C	O	0	0
			4	2	2		
4	G	1	Total	C	O	0	0
			4	2	2		
4	G	1	Total	C	O	0	0
			4	2	2		
4	G	1	Total	C	O	0	0
			4	2	2		
4	G	1	Total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

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			10	6	4		
5	B	1	Total	C	O	0	0
			10	6	4		
5	G	1	Total	C	O	0	0
			10	6	4		
5	H	1	Total	C	O	0	0
			10	6	4		
5	H	1	Total	C	O	0	0
			10	6	4		
5	H	1	Total	C	O	0	0
			10	6	4		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	O	S	0	0
			5	4	1		
6	G	1	Total	O	S	0	0
			5	4	1		
6	G	1	Total	O	S	0	0
			5	4	1		
6	G	1	Total	O	S	0	0
			5	4	1		
6	G	1	Total	O	S	0	0
			5	4	1		
6	G	1	Total	O	S	0	0
			5	4	1		
6	H	1	Total	O	S	0	0
			5	4	1		
6	H	1	Total	O	S	0	0
			5	4	1		
6	H	1	Total	O	S	0	0
			5	4	1		

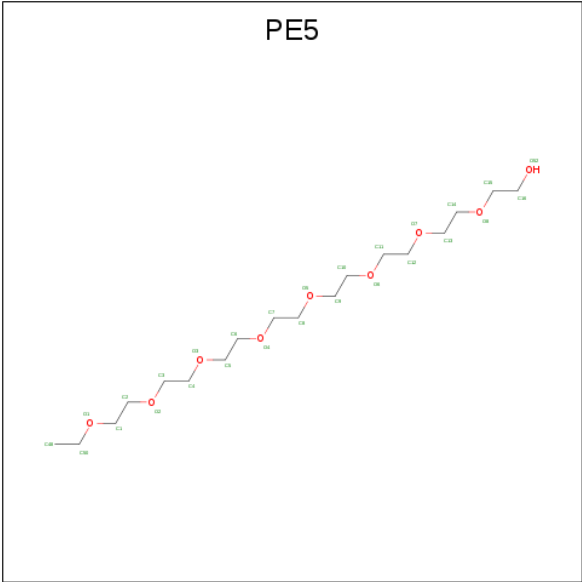
- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).





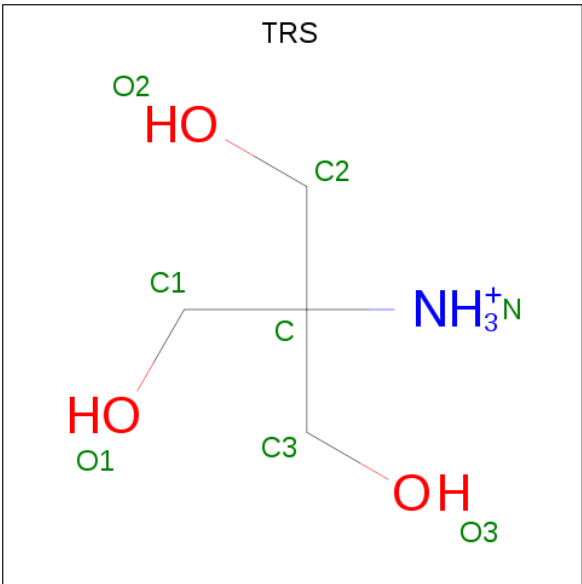
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			7	4	3		
7	B	1	Total	C	O	0	0
			7	4	3		
7	G	1	Total	C	O	0	0
			7	4	3		
7	G	1	Total	C	O	0	0
			7	4	3		
7	H	1	Total	C	O	0	0
			7	4	3		
7	H	1	Total	C	O	0	0
			7	4	3		
7	H	1	Total	C	O	0	0
			7	4	3		
7	H	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is 3,6,9,12,15,18,21,24-OCTAOXAHEXACOSAN-1-OL (three-letter code: PE5) (formula: C<sub>18</sub>H<sub>38</sub>O<sub>9</sub>).



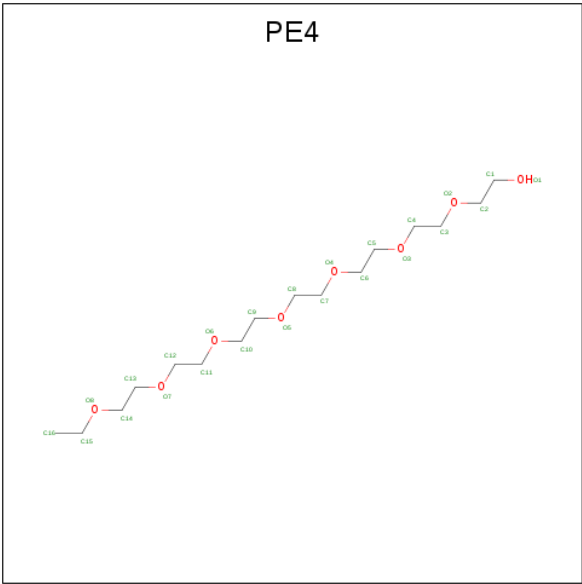
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	G	1	Total	C	O	0	0
			27	18	9		
8	H	1	Total	C	O	0	0
			27	18	9		

- Molecule 9 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	G	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 10 is 2-{2-[2-(2-{2-[2-(2-ETHOXY-ETHOXY)-ETHOXY]-ETHOXY}-ETHOXY)-ETHOXY]-ETHOXY}-ETHANOL (three-letter code: PE4) (formula: C<sub>16</sub>H<sub>34</sub>O<sub>8</sub>).

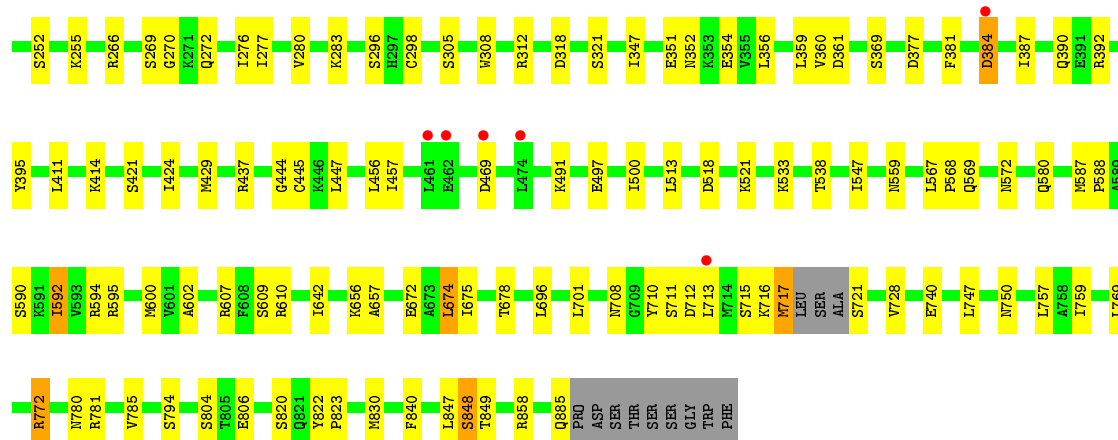


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	H	1	Total	C	O	0	0
			24	16	8		

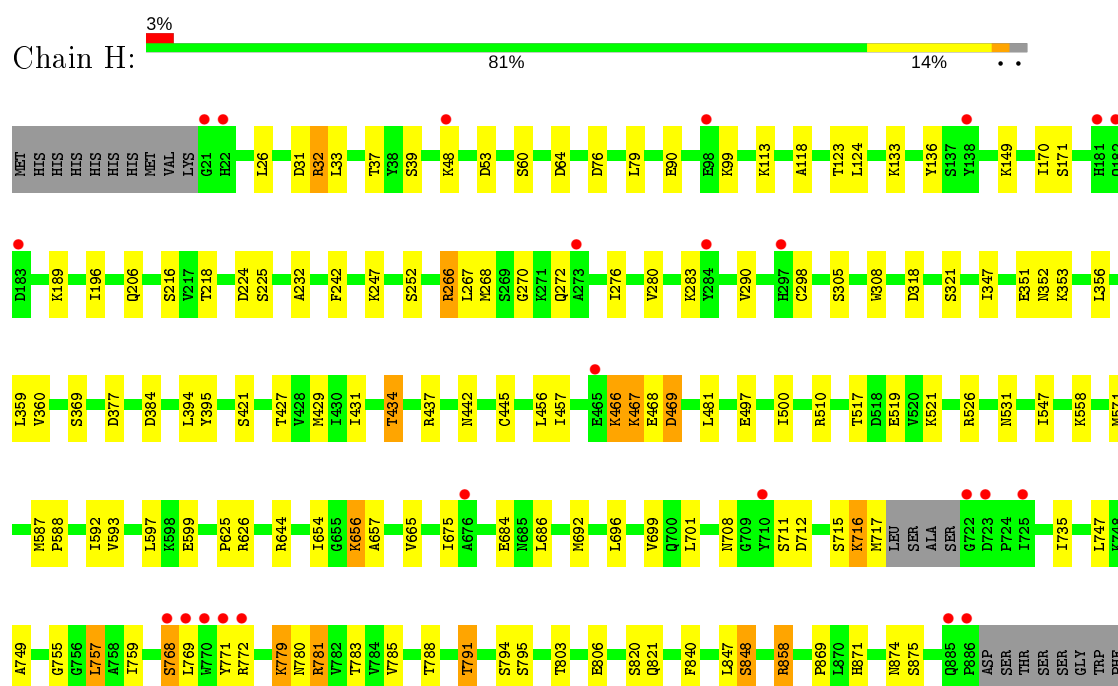
- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	22	Total	O	0	0
			22	22		
11	B	32	Total	O	0	0
			32	32		
11	G	81	Total	O	0	0
			81	81		
11	H	87	Total	O	0	0
			87	87		





• Molecule 2: Microsomal triglyceride transfer protein large subunit



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.54Å 105.59Å 112.27Å 89.81° 76.95° 74.24°	Depositor
Resolution (Å)	49.35 – 2.50 49.35 – 2.50	Depositor EDS
% Data completeness (in resolution range)	93.8 (49.35-2.50) 93.8 (49.35-2.50)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.07 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.204 , 0.257 0.204 , 0.257	Depositor DCC
$R_{free}$ test set	5414 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	65.6	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 50.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	21452	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.33% 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: PGE, PE4, PE5, CA, EDO, SO4, TRS, PEG

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	0.34	0/3756	0.68	0/5071
1	B	0.35	0/3764	0.68	0/5082
2	G	0.40	0/6897	0.71	0/9297
2	H	0.39	0/6868	0.72	0/9258
All	All	0.38	0/21285	0.70	0/28708

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	3668	0	3593	35	0
1	B	3676	0	3588	41	0
2	G	6787	0	6921	89	0
2	H	6757	0	6899	90	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	20	0	30	4	0
4	B	12	0	18	0	0
4	G	24	0	36	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	H	20	0	30	0	0
5	B	20	0	28	1	0
5	G	10	0	14	5	0
5	H	30	0	42	2	0
6	B	5	0	0	0	0
6	G	35	0	0	2	0
6	H	15	0	0	1	0
7	B	14	0	20	0	0
7	G	14	0	20	0	0
7	H	35	0	50	7	0
8	G	27	0	38	1	0
8	H	27	0	38	6	0
9	G	8	0	12	1	0
10	H	24	0	34	3	0
11	A	22	0	0	0	0
11	B	32	0	0	1	0
11	G	81	0	0	4	0
11	H	87	0	0	3	0
All	All	21452	0	21411	252	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 6.

The worst 5 of 252 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:G:717:MET:CE	2:G:721:SER:CB	2.12	1.28
2:G:717:MET:CE	2:G:721:SER:HB3	1.69	1.19
2:H:79:LEU:CD2	2:H:276:ILE:HD11	1.72	1.18
2:G:717:MET:HE2	2:G:721:SER:CB	1.73	1.15
2:G:395:TYR:CE2	5:G:902:PGE:H22	1.92	1.04

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	460/492 (94%)	436 (95%)	22 (5%)	2 (0%)	34	54
1	B	461/492 (94%)	440 (95%)	19 (4%)	2 (0%)	34	54
2	G	863/884 (98%)	831 (96%)	31 (4%)	1 (0%)	51	73
2	H	860/884 (97%)	828 (96%)	28 (3%)	4 (0%)	29	48
All	All	2644/2752 (96%)	2535 (96%)	100 (4%)	9 (0%)	41	61

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	23	GLU
1	A	166	SER
1	B	23	GLU
1	B	166	SER
2	H	716	LYS

### 5.3.2 Protein sidechains ⓘ

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	391/419 (93%)	368 (94%)	23 (6%)	19	37
1	B	393/419 (94%)	376 (96%)	17 (4%)	29	53
2	G	753/768 (98%)	708 (94%)	45 (6%)	19	37
2	H	750/768 (98%)	694 (92%)	56 (8%)	13	26
All	All	2287/2374 (96%)	2146 (94%)	141 (6%)	18	35

5 of 141 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	G	513	LEU
2	G	772	ARG
2	H	757	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	G	533	LYS
2	G	710	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
2	G	334	GLN
2	G	569	GLN
2	H	821	GLN
2	G	390	GLN
2	G	559	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 51 ligands modelled in this entry, 2 are monoatomic - leaving 49 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
7	PEG	H	918	-	6,6,6	0.16	0	5,5,5	0.04	0
4	EDO	G	912	-	3,3,3	0.07	0	2,2,2	0.16	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EDO	A	606	-	3,3,3	0.09	0	2,2,2	0.19	0
5	PGE	H	903	-	9,9,9	0.18	0	8,8,8	0.27	0
5	PGE	G	902	-	9,9,9	0.23	0	8,8,8	0.22	0
4	EDO	H	909	-	3,3,3	0.13	0	2,2,2	0.35	0
6	SO4	G	908	-	4,4,4	0.51	0	6,6,6	0.27	0
10	PE4	H	902	-	23,23,23	0.23	0	22,22,22	0.16	0
7	PEG	H	916	-	6,6,6	0.14	0	5,5,5	0.08	0
6	SO4	H	908	-	4,4,4	0.35	0	6,6,6	0.08	0
6	SO4	G	909	-	4,4,4	0.41	0	6,6,6	0.10	0
4	EDO	G	913	-	3,3,3	0.10	0	2,2,2	0.13	0
4	EDO	A	604	-	3,3,3	0.22	0	2,2,2	0.43	0
4	EDO	B	605	-	3,3,3	0.25	0	2,2,2	0.41	0
8	PE5	H	901	-	26,26,26	0.29	0	25,25,25	0.14	0
4	EDO	A	602	-	3,3,3	0.06	0	2,2,2	0.07	0
6	SO4	B	604	-	4,4,4	0.37	0	6,6,6	0.13	0
7	PEG	H	917	-	6,6,6	0.29	0	5,5,5	0.24	0
4	EDO	H	910	-	3,3,3	0.04	0	2,2,2	0.21	0
7	PEG	B	609	-	6,6,6	0.19	0	5,5,5	0.14	0
4	EDO	G	915	-	3,3,3	0.04	0	2,2,2	0.16	0
5	PGE	H	904	-	9,9,9	0.18	0	8,8,8	0.17	0
4	EDO	G	914	-	3,3,3	0.13	0	2,2,2	0.18	0
4	EDO	G	911	-	3,3,3	0.26	0	2,2,2	0.45	0
6	SO4	G	907	-	4,4,4	0.33	0	6,6,6	0.06	0
4	EDO	H	911	-	3,3,3	0.09	0	2,2,2	0.25	0
4	EDO	B	607	-	3,3,3	0.08	0	2,2,2	0.41	0
4	EDO	H	912	-	3,3,3	0.12	0	2,2,2	0.12	0
4	EDO	A	603	-	3,3,3	0.30	0	2,2,2	0.92	0
6	SO4	H	907	-	4,4,4	0.37	0	6,6,6	0.22	0
6	SO4	G	906	-	4,4,4	0.29	0	6,6,6	0.21	0
4	EDO	B	606	-	3,3,3	0.18	0	2,2,2	0.56	0
9	TRS	G	903	-	7,7,7	0.13	0	9,9,9	0.36	0
7	PEG	G	918	-	6,6,6	0.16	0	5,5,5	0.07	0
7	PEG	H	915	-	6,6,6	0.13	0	5,5,5	0.09	0
4	EDO	A	605	-	3,3,3	0.07	0	2,2,2	0.10	0
7	PEG	H	914	-	6,6,6	0.10	0	5,5,5	0.13	0
5	PGE	B	603	-	9,9,9	0.16	0	8,8,8	0.14	0
6	SO4	G	904	-	4,4,4	0.32	0	6,6,6	0.08	0
4	EDO	G	916	-	3,3,3	0.29	0	2,2,2	0.66	0
6	SO4	G	905	-	4,4,4	0.43	0	6,6,6	0.15	0
6	SO4	H	906	-	4,4,4	0.45	0	6,6,6	0.32	0
7	PEG	B	608	-	6,6,6	0.17	0	5,5,5	0.06	0
5	PGE	B	602	-	9,9,9	0.20	0	8,8,8	0.13	0
8	PE5	G	901	-	26,26,26	0.34	0	25,25,25	0.16	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	SO4	G	910	-	4,4,4	0.36	0	6,6,6	0.06	0
5	PGE	H	905	-	9,9,9	0.23	0	8,8,8	0.13	0
4	EDO	H	913	-	3,3,3	0.10	0	2,2,2	0.20	0
7	PEG	G	917	-	6,6,6	0.21	0	5,5,5	0.14	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
7	PEG	H	918	-	-	4/4/4/4	-
4	EDO	G	912	-	-	0/1/1/1	-
4	EDO	A	606	-	-	0/1/1/1	-
5	PGE	H	903	-	-	5/7/7/7	-
5	PGE	G	902	-	-	5/7/7/7	-
4	EDO	H	909	-	-	1/1/1/1	-
10	PE4	H	902	-	-	14/21/21/21	-
7	PEG	H	916	-	-	2/4/4/4	-
4	EDO	G	913	-	-	1/1/1/1	-
4	EDO	A	604	-	-	0/1/1/1	-
4	EDO	B	605	-	-	0/1/1/1	-
8	PE5	H	901	-	-	15/24/24/24	-
4	EDO	A	602	-	-	1/1/1/1	-
7	PEG	H	917	-	-	2/4/4/4	-
4	EDO	H	910	-	-	1/1/1/1	-
7	PEG	B	609	-	-	2/4/4/4	-
4	EDO	G	915	-	-	0/1/1/1	-
5	PGE	H	904	-	-	5/7/7/7	-
4	EDO	G	914	-	-	1/1/1/1	-
4	EDO	G	911	-	-	1/1/1/1	-
4	EDO	H	911	-	-	1/1/1/1	-
4	EDO	B	607	-	-	1/1/1/1	-
4	EDO	H	912	-	-	1/1/1/1	-
4	EDO	A	603	-	-	0/1/1/1	-
4	EDO	B	606	-	-	1/1/1/1	-
7	PEG	G	918	-	-	0/4/4/4	-
7	PEG	H	915	-	-	1/4/4/4	-
4	EDO	A	605	-	-	1/1/1/1	-
7	PEG	H	914	-	-	2/4/4/4	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PGE	B	603	-	-	7/7/7/7	-
4	EDO	G	916	-	-	1/1/1/1	-
9	TRS	G	903	-	-	0/9/9/9	-
7	PEG	B	608	-	-	2/4/4/4	-
5	PGE	B	602	-	-	5/7/7/7	-
8	PE5	G	901	-	-	15/24/24/24	-
5	PGE	H	905	-	-	3/7/7/7	-
4	EDO	H	913	-	-	0/1/1/1	-
7	PEG	G	917	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 104 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
10	H	902	PE4	C3-C4-O3-C5
5	G	902	PGE	C3-C4-O3-C5
7	G	917	PEG	O1-C1-C2-O2
5	G	902	PGE	O2-C3-C4-O3
8	H	901	PE5	O2-C3-C4-O3

There are no ring outliers.

17 monomers are involved in 34 short contacts:

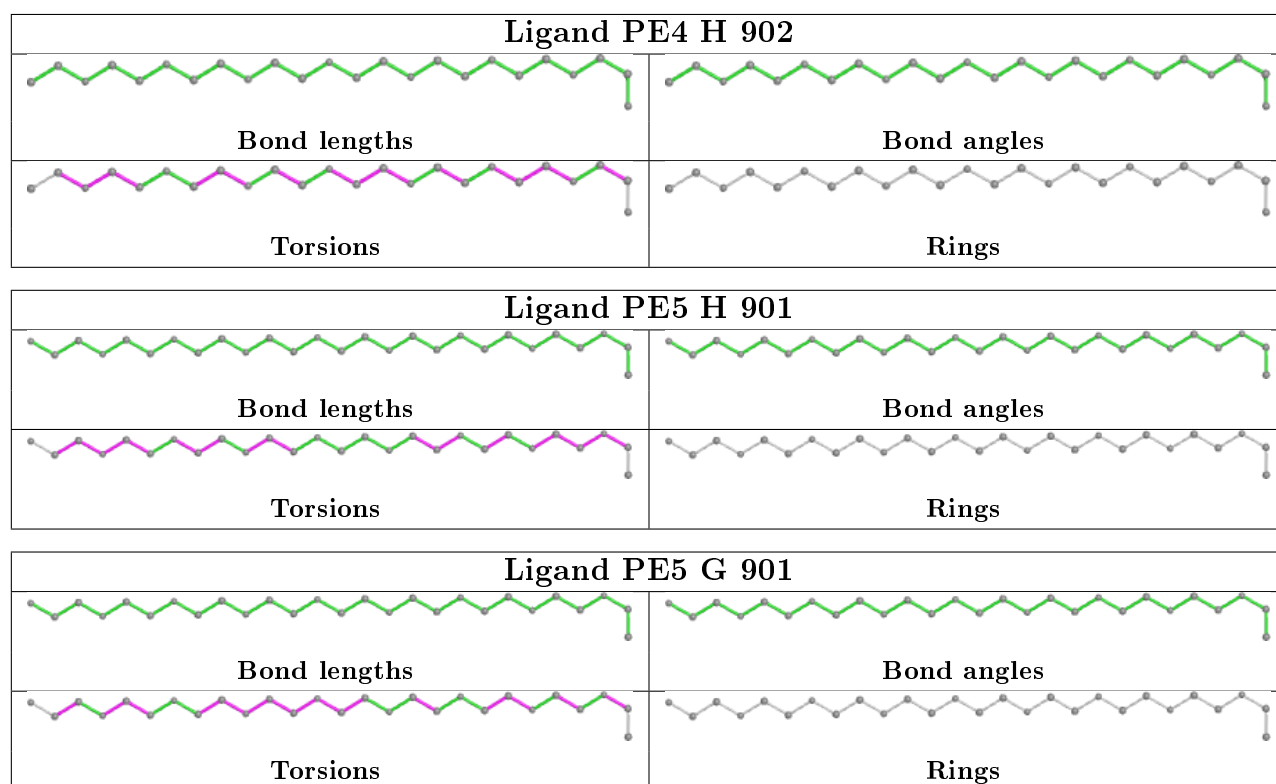
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	H	903	PGE	1	0
5	G	902	PGE	5	0
10	H	902	PE4	3	0
4	A	604	EDO	1	0
8	H	901	PE5	6	0
4	A	602	EDO	1	0
7	H	917	PEG	5	0
4	G	915	EDO	1	0
6	G	907	SO4	1	0
4	A	603	EDO	2	0
6	G	906	SO4	1	0
9	G	903	TRS	1	0
7	H	914	PEG	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	603	PGE	1	0
6	H	906	SO4	1	0
8	G	901	PE5	1	0
5	H	905	PGE	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 ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	461/492 (93%)	0.58	68 (14%) 2 2	38, 76, 149, 188	0
1	B	462/492 (93%)	0.16	23 (4%) 28 30	39, 72, 120, 167	0
2	G	863/884 (97%)	0.03	11 (1%) 77 79	34, 61, 107, 151	0
2	H	862/884 (97%)	0.18	24 (2%) 53 56	35, 64, 122, 167	0
All	All	2648/2752 (96%)	0.20	126 (4%) 30 32	34, 66, 127, 188	0

The worst 5 of 126 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	77	ILE	11.4
1	B	72	ALA	7.6
1	A	79	LEU	7.2
1	A	72	ALA	6.7
1	A	111	ALA	6.7

### 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 carbohydrates 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
4	EDO	G	913	4/4	0.54	0.28	90,96,103,106	0
4	EDO	G	912	4/4	0.55	0.21	72,78,81,83	4
7	PEG	B	608	7/7	0.55	0.34	65,74,81,82	7
8	PE5	G	901	27/27	0.55	0.75	85,106,117,126	27
4	EDO	H	910	4/4	0.62	0.21	94,98,102,104	0
4	EDO	G	914	4/4	0.63	0.39	101,104,110,113	0
8	PE5	H	901	27/27	0.64	0.59	70,98,130,142	27
7	PEG	H	916	7/7	0.66	0.51	89,101,106,106	7
4	EDO	H	913	4/4	0.67	0.19	78,81,82,85	4
7	PEG	H	918	7/7	0.69	0.40	71,78,87,95	7
4	EDO	H	911	4/4	0.70	0.34	81,88,93,94	0
4	EDO	A	604	4/4	0.71	0.20	71,84,84,92	0
4	EDO	H	909	4/4	0.72	0.30	75,82,91,91	0
7	PEG	B	609	7/7	0.72	0.77	60,69,80,81	7
4	EDO	B	606	4/4	0.72	0.28	88,90,96,96	0
10	PE4	H	902	24/24	0.73	0.24	84,109,117,122	24
4	EDO	H	912	4/4	0.73	0.34	98,103,111,115	0
4	EDO	A	606	4/4	0.76	0.21	96,106,110,114	0
5	PGE	B	602	10/10	0.76	0.59	67,76,90,93	10
5	PGE	B	603	10/10	0.79	0.43	92,98,103,105	10
7	PEG	H	917	7/7	0.80	0.24	66,71,78,78	7
7	PEG	G	918	7/7	0.82	0.31	62,79,88,89	7
4	EDO	G	911	4/4	0.82	0.14	64,79,85,91	0
4	EDO	A	602	4/4	0.82	0.17	68,71,73,75	0
5	PGE	H	904	10/10	0.83	0.32	67,78,96,99	10
5	PGE	H	903	10/10	0.83	0.38	64,73,81,85	10
5	PGE	G	902	10/10	0.83	0.42	65,69,88,88	10
7	PEG	H	915	7/7	0.84	0.26	76,77,81,82	7
4	EDO	A	605	4/4	0.84	0.21	62,72,73,76	0
4	EDO	B	605	4/4	0.85	0.17	79,89,90,92	0
4	EDO	A	603	4/4	0.85	0.54	64,66,67,68	0
6	SO4	G	906	5/5	0.85	0.25	64,65,74,75	5
4	EDO	G	916	4/4	0.86	0.22	64,69,71,75	0
7	PEG	H	914	7/7	0.87	0.23	77,89,90,91	0
5	PGE	H	905	10/10	0.87	0.33	83,89,102,103	0
6	SO4	G	909	5/5	0.87	0.36	78,81,86,89	5
6	SO4	G	910	5/5	0.88	0.32	87,87,90,91	5
9	TRS	G	903	8/8	0.88	0.27	51,62,66,74	0
4	EDO	G	915	4/4	0.88	0.25	57,60,63,65	4
6	SO4	B	604	5/5	0.89	0.29	63,65,73,76	5
4	EDO	B	607	4/4	0.90	0.18	69,72,76,78	4
6	SO4	G	907	5/5	0.91	0.18	76,78,85,87	5
7	PEG	G	917	7/7	0.91	0.38	67,83,95,102	0

Continued on next page...

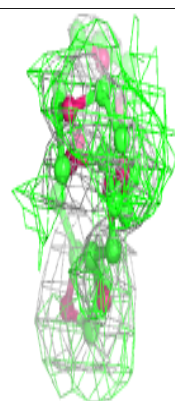
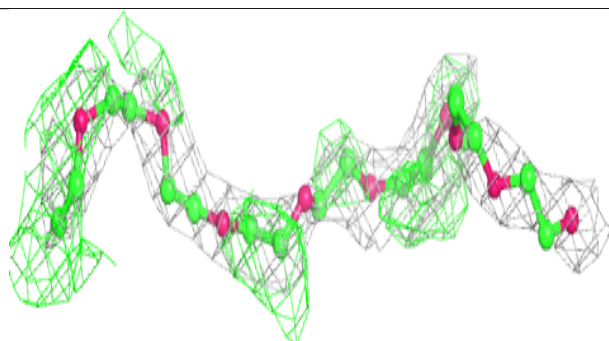
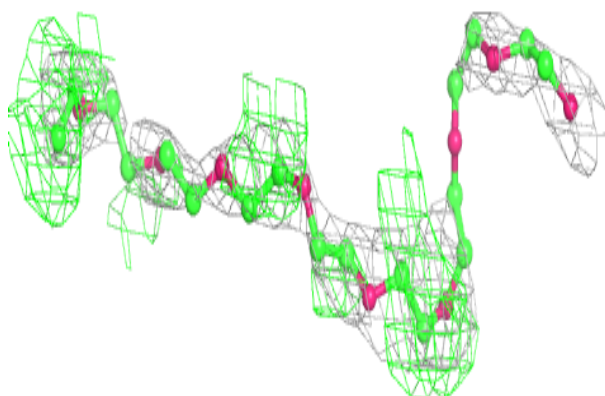
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	SO4	G	905	5/5	0.92	0.29	66,68,72,74	5
6	SO4	G	904	5/5	0.93	0.23	77,78,83,88	5
6	SO4	H	908	5/5	0.95	0.25	66,69,74,76	5
6	SO4	G	908	5/5	0.95	0.21	47,48,54,57	0
3	CA	A	601	1/1	0.96	0.06	57,57,57,57	0
3	CA	B	601	1/1	0.96	0.09	60,60,60,60	0
6	SO4	H	907	5/5	0.97	0.19	63,65,72,72	5
6	SO4	H	906	5/5	0.98	0.12	66,68,80,80	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.

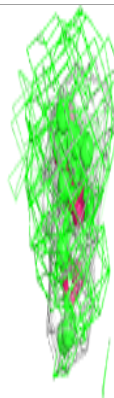
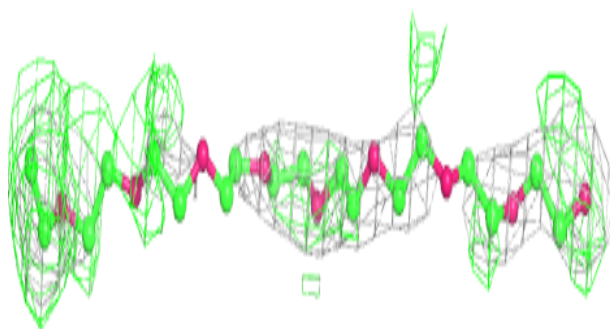
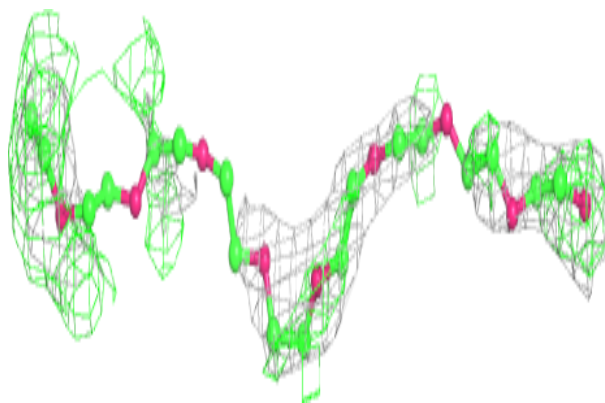
**Electron density around PE5 G 901:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)

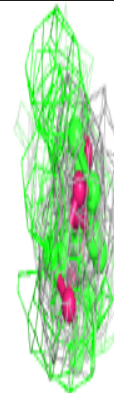
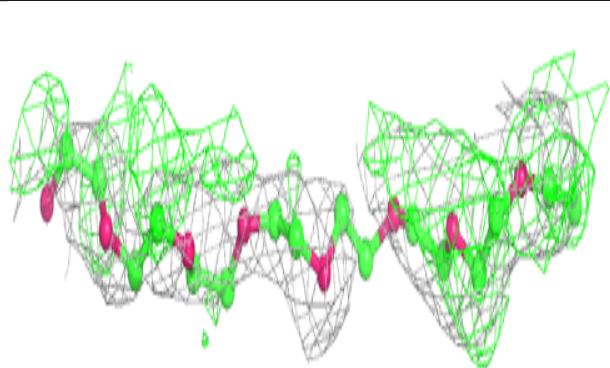
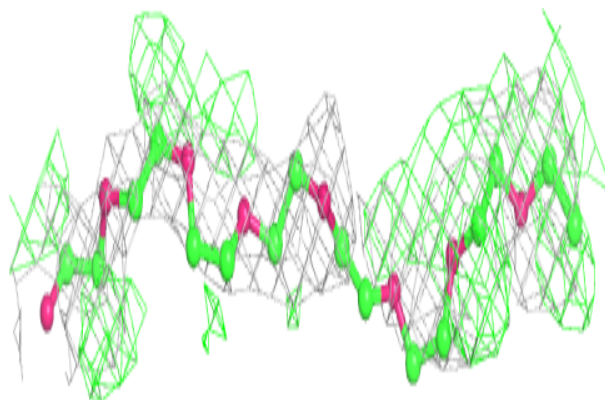


**Electron density around PE5 H 901:**

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

**Electron density around PE4 H 902:**

$2mF_o-DF_c$  (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.