



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 27, 2022 – 02:05 PM EDT

PDB ID : 7RE7  
Title : TCR mimic antibody (Fab fragment) in complex with AFP/HLA-A\*02  
Authors : Dasgupta, M.; Baker, B.M.  
Deposited on : 2021-07-12  
Resolution : 2.55 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 types of validation reports are described at <http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.29
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)	:	2.29

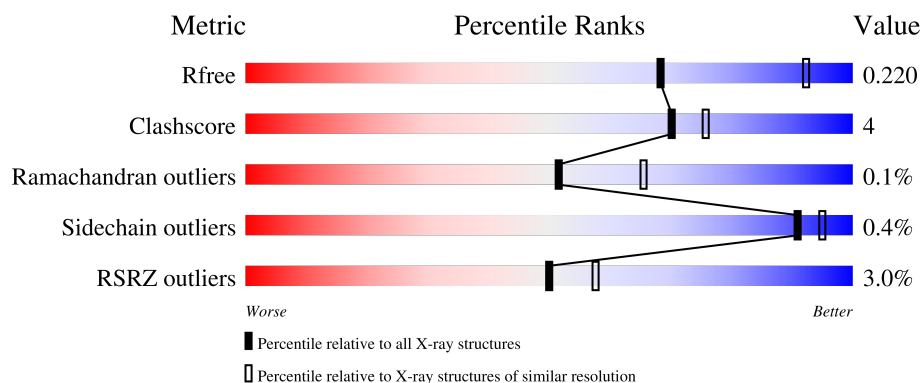
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 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.





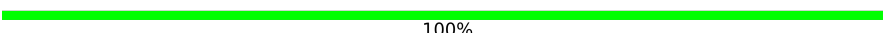

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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	H	224	<div> <div>83%</div> <div>12%</div> <div>5%</div> </div>
1	I	224	<div> <div>89%</div> <div>8%</div> <div>.</div> </div>
2	J	216	<div> <div>4%</div> <div>93%</div> <div>5%</div> <div>.</div> </div>
2	L	216	<div> <div>14%</div> <div>90%</div> <div>8%</div> <div>.</div> </div>
3	A	295	<div> <div>85%</div> <div>7%</div> <div>6%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	D	295	 86% 7% 6%
4	B	100	 90% 10%
4	E	100	 96% . .
5	C	9	 100%
5	F	9	 89% 11%

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
6	EPE	B	202	-	-	X	-

## 2 Entry composition

There are 13 unique types of molecules in this entry. The entry contains 13099 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 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	218	Total	C	N	O	S	0	0	0
			1646	1046	269	324	7			
1	H	212	Total	C	N	O	S	0	0	0
			1605	1022	261	315	7			

- Molecule 2 is a protein called Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	J	211	Total	C	N	O	S	0	0	0
			1569	977	264	323	5			
2	L	211	Total	C	N	O	S	0	0	0
			1569	977	264	323	5			

- Molecule 3 is a protein called MHC class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	276	Total	C	N	O	S	0	0	0
			2253	1408	410	426	9			
3	D	277	Total	C	N	O	S	0	0	0
			2257	1410	411	427	9			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	277	GLY	SER	conflict	UNP Q861F7
A	279	GLY	-	expression tag	UNP Q861F7
A	280	GLY	-	expression tag	UNP Q861F7
A	281	GLY	-	expression tag	UNP Q861F7
A	282	LEU	-	expression tag	UNP Q861F7
A	283	ASN	-	expression tag	UNP Q861F7
A	284	ASP	-	expression tag	UNP Q861F7
A	285	ILE	-	expression tag	UNP Q861F7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	286	PHE	-	expression tag	UNP Q861F7
A	287	GLU	-	expression tag	UNP Q861F7
A	288	ALA	-	expression tag	UNP Q861F7
A	289	GLN	-	expression tag	UNP Q861F7
A	290	LYS	-	expression tag	UNP Q861F7
A	291	ILE	-	expression tag	UNP Q861F7
A	292	GLY	-	expression tag	UNP Q861F7
A	293	TRP	-	expression tag	UNP Q861F7
A	294	HIS	-	expression tag	UNP Q861F7
A	295	GLU	-	expression tag	UNP Q861F7
D	277	GLY	SER	conflict	UNP Q861F7
D	279	GLY	-	expression tag	UNP Q861F7
D	280	GLY	-	expression tag	UNP Q861F7
D	281	GLY	-	expression tag	UNP Q861F7
D	282	LEU	-	expression tag	UNP Q861F7
D	283	ASN	-	expression tag	UNP Q861F7
D	284	ASP	-	expression tag	UNP Q861F7
D	285	ILE	-	expression tag	UNP Q861F7
D	286	PHE	-	expression tag	UNP Q861F7
D	287	GLU	-	expression tag	UNP Q861F7
D	288	ALA	-	expression tag	UNP Q861F7
D	289	GLN	-	expression tag	UNP Q861F7
D	290	LYS	-	expression tag	UNP Q861F7
D	291	ILE	-	expression tag	UNP Q861F7
D	292	GLY	-	expression tag	UNP Q861F7
D	293	TRP	-	expression tag	UNP Q861F7
D	294	HIS	-	expression tag	UNP Q861F7
D	295	GLU	-	expression tag	UNP Q861F7

- Molecule 4 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	100	Total	C	N	O	S	0	0	0
			837	533	141	159	4			
4	E	100	Total	C	N	O	S	0	0	0
			837	533	141	159	4			

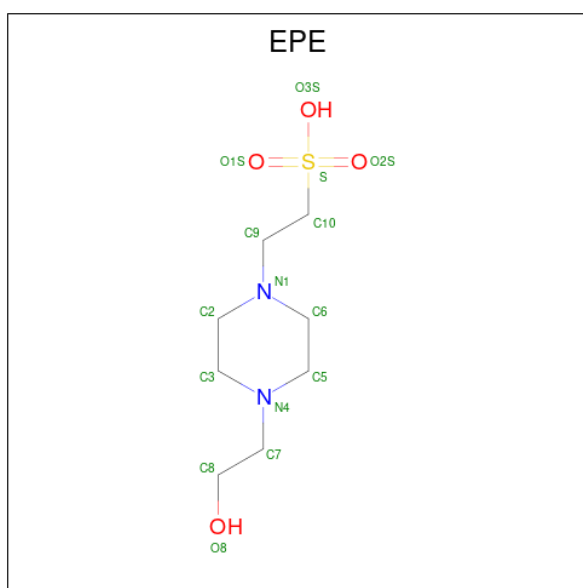
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	initiating methionine	UNP P61769
E	1	MET	-	initiating methionine	UNP P61769

- Molecule 5 is a protein called PHE-MET-ASN-LYS-PHE-ILE-TYR-GLU-ILE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	9	Total	C	N	O	S	0	0	0
			85	59	11	14	1			
5	F	9	Total	C	N	O	S	0	0	0
			85	59	11	14	1			

- Molecule 6 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula:  $C_8H_{18}N_2O_4S$ ).



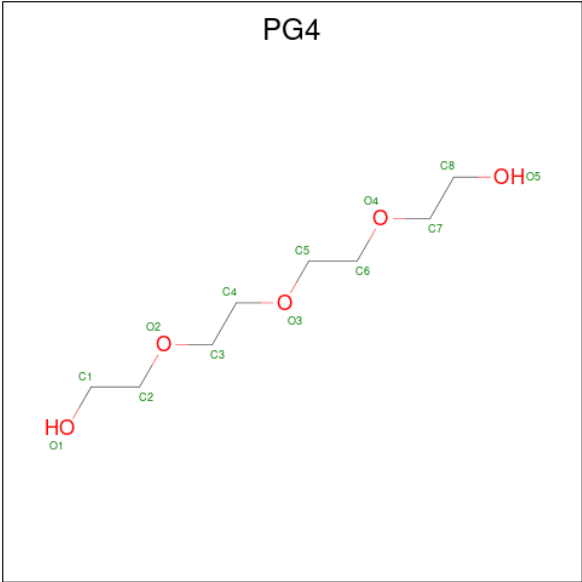
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	I	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
6	H	1	Total	C	N	O	S	0	0
			15	8	2	4	1		
6	B	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



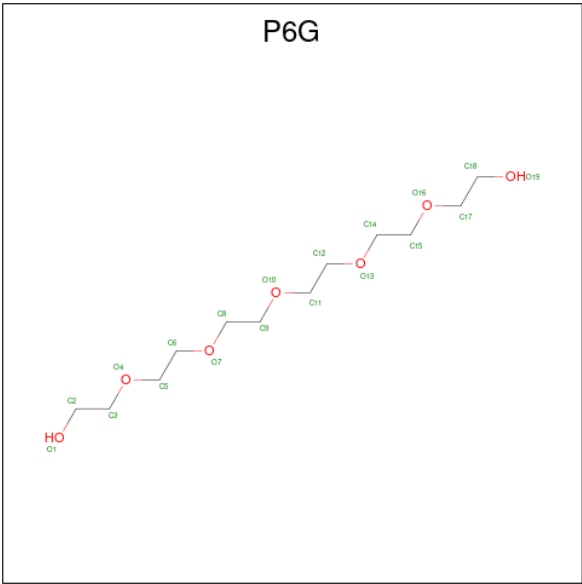
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	I	1	Total	C	O	0	0
			6	3	3		
7	H	1	Total	C	O	0	0
			6	3	3		
7	H	1	Total	C	O	0	0
			6	3	3		
7	A	1	Total	C	O	0	0
			6	3	3		
7	D	1	Total	C	O	0	0
			6	3	3		
7	E	1	Total	C	O	0	0
			6	3	3		
7	E	1	Total	C	O	0	0
			6	3	3		

- Molecule 8 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			13	8	5		
8	A	1	Total	C	O	0	0
			13	8	5		
8	D	1	Total	C	O	0	0
			13	8	5		
8	E	1	Total	C	O	0	0
			13	8	5		

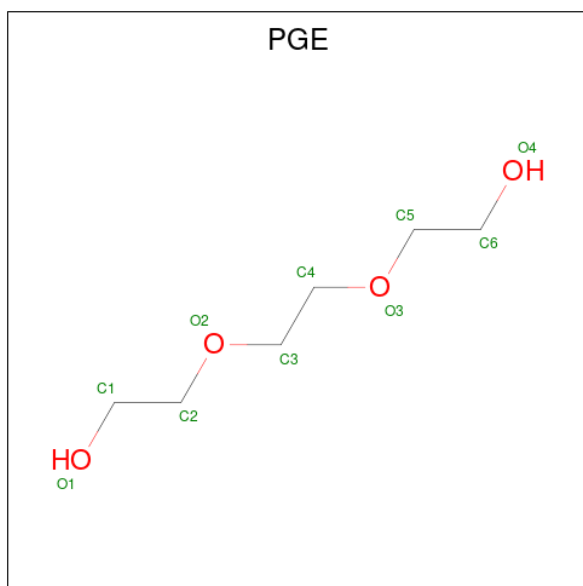
- Molecule 9 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: C<sub>12</sub>H<sub>26</sub>O<sub>7</sub>).





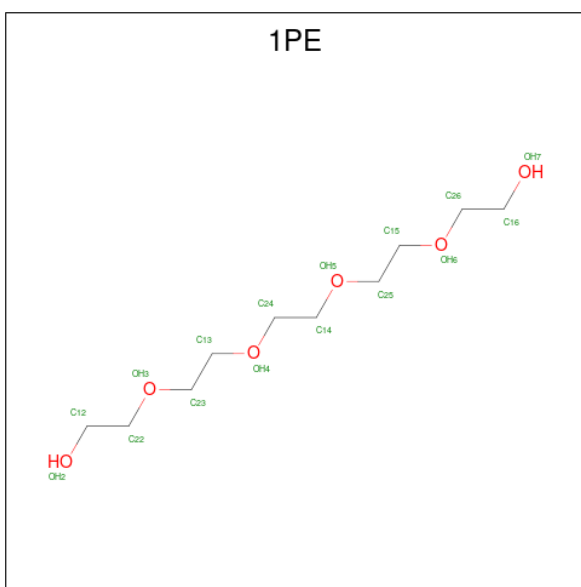
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	B	1	Total	C	O	0	0
			19	12	7		

- Molecule 10 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula:  $C_6H_{14}O_4$ ).



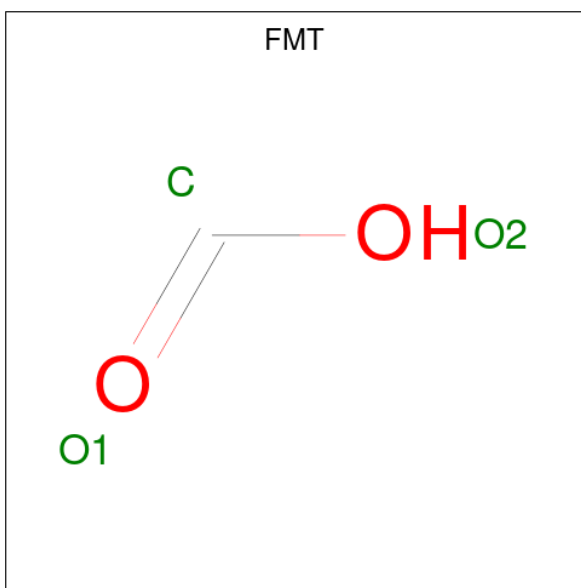
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	B	1	Total	C	O	0	0
			10	6	4		
10	E	1	Total	C	O	0	0
			10	6	4		

- Molecule 11 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula:  $C_{10}H_{22}O_6$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	D	1	Total	C	O	0	0
			16	10	6		

- Molecule 12 is FORMIC ACID (three-letter code: FMT) (formula:  $\text{CH}_2\text{O}_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
12	D	1	Total	C	O	0	0
			3	1	2		

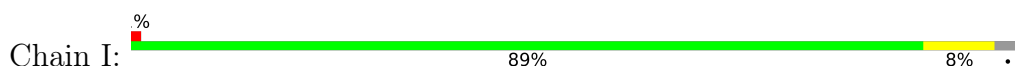
- Molecule 13 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	I	20	Total 20	O 20	0	0
13	J	16	Total 16	O 16	0	0
13	H	20	Total 20	O 20	0	0
13	L	13	Total 13	O 13	0	0
13	A	31	Total 31	O 31	0	0
13	B	10	Total 10	O 10	0	0
13	D	27	Total 27	O 27	0	0
13	E	16	Total 16	O 16	0	0
13	C	3	Total 3	O 3	0	0
13	F	3	Total 3	O 3	0	0

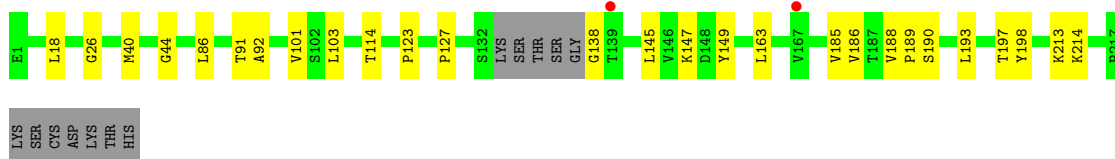
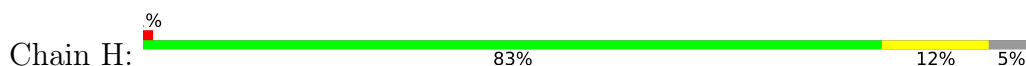
### 3 Residue-property plots [i](#)

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: Fab heavy chain



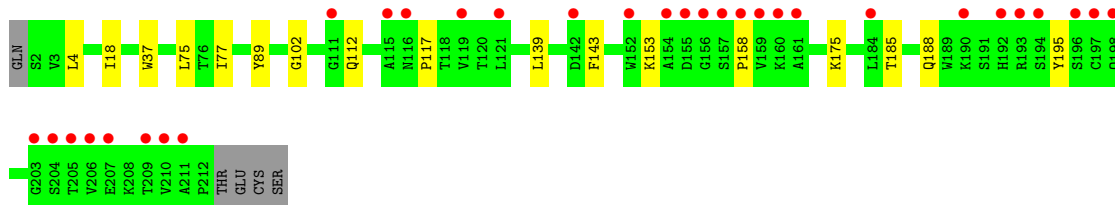
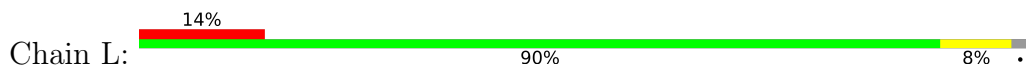
- Molecule 1: Fab heavy chain



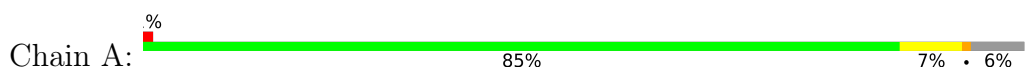
- Molecule 2: Fab light chain

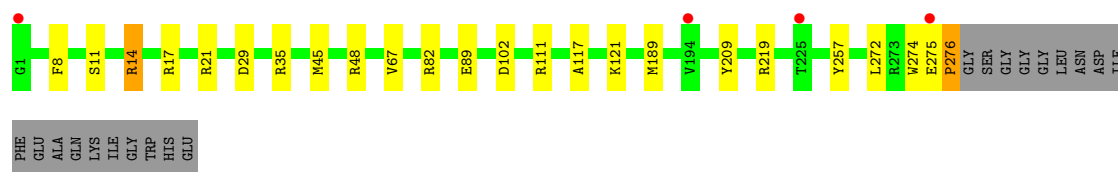


- Molecule 2: Fab light chain



- Molecule 3: MHC class I antigen





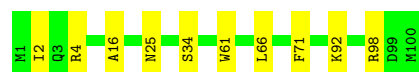
- Molecule 3: MHC class I antigen

Chain D: 86% 7% 6%



- Molecule 4: Beta-2-microglobulin

Chain B: 90% 10%



- Molecule 4: Beta-2-microglobulin

Chain E: 96% 4%



- Molecule 5: PHE-MET-ASN-LYS-PHE-ILE-TYR-GLU-ILE

Chain C: 100%

There are no outlier residues recorded for this chain.

- Molecule 5: PHE-MET-ASN-LYS-PHE-ILE-TYR-GLU-ILE

Chain F: 89% 11%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.71Å 200.91Å 244.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.20 – 2.55 49.20 – 2.55	Depositor EDS
% Data completeness (in resolution range)	96.4 (49.20-2.55) 95.7 (49.20-2.55)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 2.54Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, $R_{free}$	0.180 , 0.220 0.186 , 0.220	Depositor DCC
$R_{free}$ test set	8113 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.4	Xtriage
Anisotropy	0.821	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 34.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13099	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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: PGE, EPE, P6G, 1PE, PG4, GOL, FMT

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	H	0.33	0/1649	0.47	0/2254
1	I	0.49	2/1691 (0.1%)	0.53	2/2310 (0.1%)
2	J	0.33	0/1606	0.46	0/2190
2	L	0.34	0/1606	0.46	0/2190
3	A	0.30	1/2319 (0.0%)	0.48	1/3149 (0.0%)
3	D	0.31	0/2323	0.47	0/3154
4	B	0.24	0/860	0.44	0/1162
4	E	0.27	0/860	0.45	0/1162
5	C	0.28	0/87	0.39	0/113
5	F	0.27	0/87	0.36	0/113
All	All	0.34	3/13088 (0.0%)	0.48	3/17797 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	I	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	217	PRO	N-CA	13.41	1.70	1.47
3	A	276	PRO	N-CD	-6.39	1.39	1.47
1	I	216	GLU	C-N	5.24	1.44	1.34

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	217	PRO	N-CA-C	-5.87	96.84	112.10
1	I	217	PRO	CA-N-CD	-5.46	103.86	111.50
3	A	276	PRO	CA-N-CD	5.12	118.86	111.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	217	PRO	Peptide

## 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	H	1605	0	1559	15	0
1	I	1646	0	1606	13	0
2	J	1569	0	1515	8	0
2	L	1569	0	1515	11	0
3	A	2253	0	2103	25	0
3	D	2257	0	2106	15	0
4	B	837	0	803	6	0
4	E	837	0	803	6	0
5	C	85	0	85	0	0
5	F	85	0	85	1	0
6	B	15	0	17	8	0
6	H	15	0	18	2	0
6	I	15	0	18	0	0
7	A	6	0	8	0	0
7	D	6	0	8	0	0
7	E	12	0	16	2	0
7	H	12	0	16	0	0
7	I	6	0	8	0	0
8	A	26	0	36	1	0
8	D	13	0	18	0	0
8	E	13	0	18	0	0
9	B	19	0	26	2	0
10	B	10	0	14	1	0
10	E	10	0	14	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	D	16	0	22	2	0
12	D	3	0	1	0	0
13	A	31	0	0	0	0
13	B	10	0	0	0	0
13	C	3	0	0	0	0
13	D	27	0	0	0	0
13	E	16	0	0	0	0
13	F	3	0	0	0	0
13	H	20	0	0	0	0
13	I	20	0	0	0	0
13	J	16	0	0	0	0
13	L	13	0	0	0	0
All	All	13099	0	12438	92	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.

The worst 5 of 92 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:I:217:PRO:N	1:I:217:PRO:CA	1.70	1.42
3:A:275:GLU:HB2	3:A:276:PRO:HD3	1.68	0.74
4:B:92:LYS:HD3	10:B:203:PGE:H5	1.70	0.72
1:I:217:PRO:N	1:I:217:PRO:C	2.42	0.71
4:B:34:SER:OG	6:B:202:EPE:H82	1.90	0.71

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	H	208/224 (93%)	205 (99%)	2 (1%)	1 (0%)	29	40
1	I	216/224 (96%)	209 (97%)	7 (3%)	0	100	100
2	J	209/216 (97%)	206 (99%)	3 (1%)	0	100	100
2	L	209/216 (97%)	199 (95%)	10 (5%)	0	100	100
3	A	274/295 (93%)	271 (99%)	3 (1%)	0	100	100
3	D	275/295 (93%)	269 (98%)	6 (2%)	0	100	100
4	B	98/100 (98%)	97 (99%)	1 (1%)	0	100	100
4	E	98/100 (98%)	96 (98%)	2 (2%)	0	100	100
5	C	7/9 (78%)	7 (100%)	0	0	100	100
5	F	7/9 (78%)	7 (100%)	0	0	100	100
All	All	1601/1688 (95%)	1566 (98%)	34 (2%)	1 (0%)	51	65

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	101	VAL

### 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	H	182/193 (94%)	182 (100%)	0	100	100
1	I	187/193 (97%)	186 (100%)	1 (0%)	88	93
2	J	177/182 (97%)	177 (100%)	0	100	100
2	L	177/182 (97%)	177 (100%)	0	100	100
3	A	232/245 (95%)	231 (100%)	1 (0%)	91	95
3	D	232/245 (95%)	231 (100%)	1 (0%)	91	95
4	B	95/95 (100%)	94 (99%)	1 (1%)	73	83
4	E	95/95 (100%)	94 (99%)	1 (1%)	73	83
5	C	9/9 (100%)	9 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	F	9/9 (100%)	9 (100%)	0	100	100
All	All	1395/1448 (96%)	1390 (100%)	5 (0%)	91	95

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

Mol	Chain	Res	Type
1	I	200	CYS
3	A	14	ARG
4	B	71	PHE
3	D	275	GLU
4	E	98	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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](#)

19 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	GOL	E	604	-	5,5,5	0.56	0	5,5,5	0.23	0
8	PG4	A	302	-	12,12,12	0.52	0	11,11,11	0.20	0
7	GOL	H	302	-	5,5,5	0.56	0	5,5,5	0.22	0
6	EPE	I	301	-	15,15,15	1.39	3 (20%)	18,20,20	2.40	9 (50%)
10	PGE	B	203	-	9,9,9	0.30	0	8,8,8	0.30	0
9	P6G	B	201	-	18,18,18	0.53	0	17,17,17	0.24	0
6	EPE	B	202	-	15,15,15	2.42	1 (6%)	18,20,20	2.24	5 (27%)
7	GOL	E	603	-	5,5,5	0.53	0	5,5,5	0.28	0
7	GOL	I	302	-	5,5,5	0.53	0	5,5,5	0.32	0
8	PG4	D	304	-	12,12,12	0.53	0	11,11,11	0.24	0
10	PGE	E	602	-	9,9,9	0.31	0	8,8,8	0.30	0
12	FMT	D	302	-	2,2,2	0.72	0	1,1,1	0.25	0
7	GOL	H	303	-	5,5,5	0.54	0	5,5,5	0.25	0
11	1PE	D	301	-	15,15,15	0.52	0	14,14,14	0.27	0
8	PG4	A	303	-	12,12,12	0.52	0	11,11,11	0.24	0
7	GOL	A	301	-	5,5,5	0.55	0	5,5,5	0.25	0
6	EPE	H	301	-	15,15,15	1.29	3 (20%)	18,20,20	2.36	7 (38%)
8	PG4	E	601	-	12,12,12	0.54	0	11,11,11	0.26	0
7	GOL	D	303	-	5,5,5	0.55	0	5,5,5	0.27	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
6	EPE	H	301	-	-	4/9/19/19	0/1/1/1
7	GOL	I	302	-	-	2/4/4/4	-
7	GOL	E	604	-	-	2/4/4/4	-
8	PG4	A	302	-	-	1/10/10/10	-
7	GOL	H	302	-	-	0/4/4/4	-
8	PG4	E	601	-	-	4/10/10/10	-
8	PG4	D	304	-	-	4/10/10/10	-
9	P6G	B	201	-	-	6/16/16/16	-
10	PGE	E	602	-	-	3/7/7/7	-
6	EPE	I	301	-	-	3/9/19/19	0/1/1/1
7	GOL	H	303	-	-	0/4/4/4	-
11	1PE	D	301	-	-	9/13/13/13	-
10	PGE	B	203	-	-	4/7/7/7	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	PG4	A	303	-	-	5/10/10/10	-
7	GOL	D	303	-	-	0/4/4/4	-
6	EPE	B	202	-	-	4/9/19/19	0/1/1/1
7	GOL	A	301	-	-	0/4/4/4	-
7	GOL	E	603	-	-	0/4/4/4	-

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	202	EPE	C10-S	-8.90	1.64	1.77
6	I	301	EPE	C10-S	3.86	1.83	1.77
6	H	301	EPE	C10-S	3.27	1.82	1.77
6	H	301	EPE	O1S-S	2.23	1.51	1.45
6	I	301	EPE	O2S-S	2.23	1.51	1.45

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	202	EPE	O2S-S-C10	4.84	112.74	106.92
6	I	301	EPE	O1S-S-C10	4.59	112.45	106.92
6	B	202	EPE	C2-C3-N4	-4.33	101.77	110.64
6	H	301	EPE	O3S-S-O2S	-4.29	100.78	111.27
6	H	301	EPE	C2-C3-N4	3.99	118.82	110.64

There are no chirality outliers.

5 of 51 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	H	301	EPE	S-C10-C9-N1
6	B	202	EPE	C10-C9-N1-C6
7	I	302	GOL	C1-C2-C3-O3
7	E	604	GOL	C1-C2-C3-O3
8	E	601	PG4	O2-C3-C4-O3

There are no ring outliers.

7 monomers are involved in 18 short contacts:

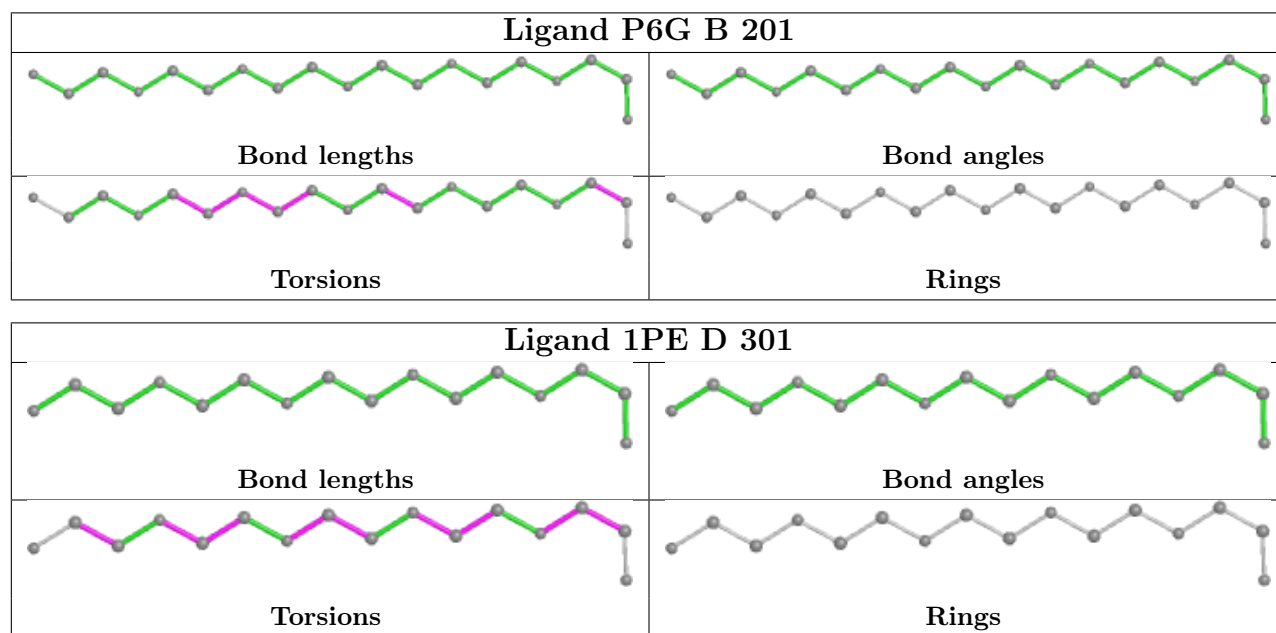
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	E	604	GOL	2	0
10	B	203	PGE	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	B	201	P6G	2	0
6	B	202	EPE	8	0
11	D	301	1PE	2	0
8	A	303	PG4	1	0
6	H	301	EPE	2	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	H	212/224 (94%)	-0.12	2 (0%) 84 88	28, 48, 91, 132	0
1	I	218/224 (97%)	0.01	2 (0%) 84 88	29, 54, 82, 121	0
2	J	211/216 (97%)	0.05	9 (4%) 35 42	32, 51, 107, 123	0
2	L	211/216 (97%)	0.64	31 (14%) 2 3	26, 64, 136, 182	0
3	A	276/295 (93%)	-0.12	4 (1%) 75 81	23, 36, 75, 139	0
3	D	277/295 (93%)	-0.17	0 100 100	27, 39, 65, 102	0
4	B	100/100 (100%)	-0.23	0 100 100	28, 46, 79, 96	0
4	E	100/100 (100%)	0.03	1 (1%) 82 86	32, 50, 84, 98	0
5	C	9/9 (100%)	-0.09	0 100 100	24, 28, 35, 43	0
5	F	9/9 (100%)	0.02	0 100 100	28, 33, 43, 43	0
All	All	1623/1688 (96%)	0.01	49 (3%) 50 57	23, 46, 101, 182	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	L	158	PRO	9.0
2	L	193	ARG	5.1
2	L	159	VAL	4.4
2	L	206	VAL	3.9
2	L	154	ALA	3.9

### 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.

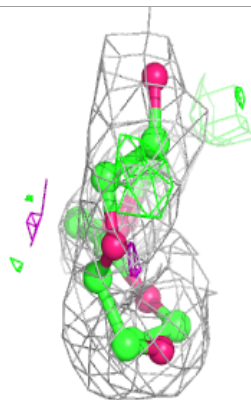
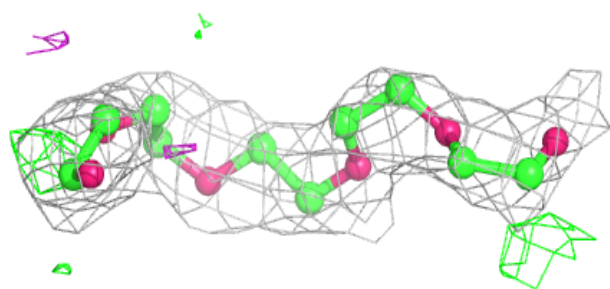
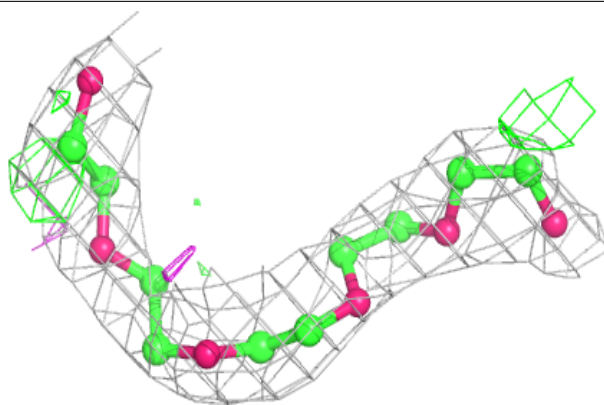
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	GOL	I	302	6/6	0.67	0.23	72,77,77,79	0
7	GOL	H	303	6/6	0.79	0.19	52,63,66,75	0
7	GOL	D	303	6/6	0.80	0.28	73,79,83,84	0
8	PG4	D	304	13/13	0.81	0.29	60,77,89,90	0
11	1PE	D	301	16/16	0.81	0.24	60,72,74,78	0
10	PGE	E	602	10/10	0.83	0.22	64,68,70,70	0
7	GOL	E	604	6/6	0.83	0.23	69,72,75,75	0
8	PG4	A	303	13/13	0.85	0.20	70,76,78,80	0
7	GOL	A	301	6/6	0.85	0.33	83,86,89,93	0
6	EPE	B	202	15/15	0.86	0.31	71,84,124,127	0
6	EPE	I	301	15/15	0.86	0.25	26,58,118,119	0
6	EPE	H	301	15/15	0.88	0.24	59,71,110,112	0
7	GOL	H	302	6/6	0.88	0.27	62,72,75,77	0
12	FMT	D	302	3/3	0.89	0.15	61,61,63,65	0
10	PGE	B	203	10/10	0.90	0.24	68,72,75,78	0
7	GOL	E	603	6/6	0.90	0.19	52,55,57,64	0
8	PG4	A	302	13/13	0.91	0.16	49,61,66,69	0
8	PG4	E	601	13/13	0.93	0.17	44,53,58,63	0
9	P6G	B	201	19/19	0.96	0.20	35,53,81,81	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.

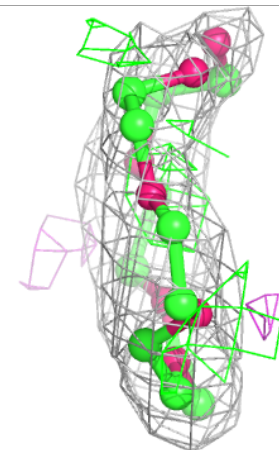
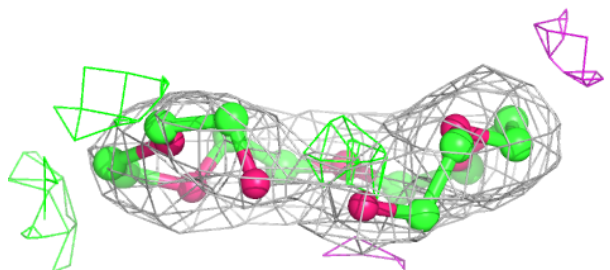
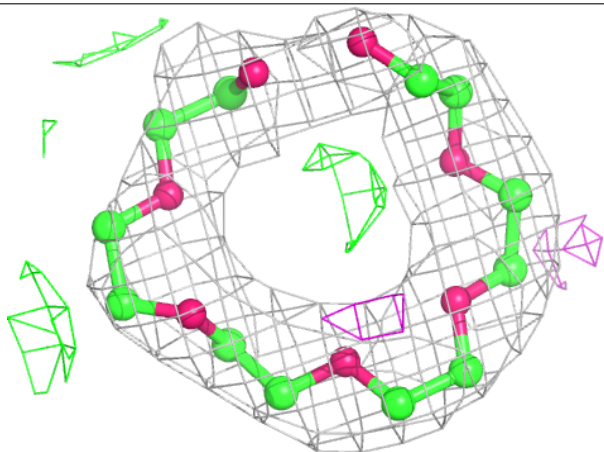


**Electron density around 1PE D 301:**

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

**Electron density around P6G B 201:**

$2mF_o-DF_c$  (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.