



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 18, 2023 – 12:09 PM EDT

PDB ID : 8EM5  
Title : Mycobacterium thermoresistibile MmpS5  
Authors : Cuthbert, B.J.; Goulding, C.W.  
Deposited on : 2022-09-26  
Resolution : 1.95 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

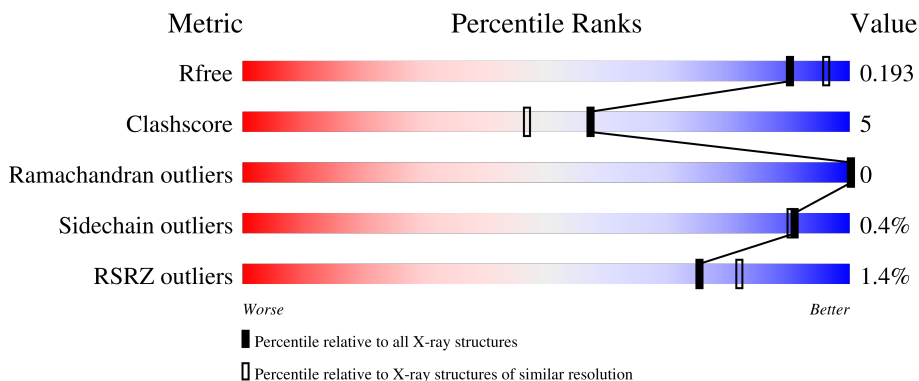
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	103	
1	B	103	
1	C	103	
1	D	103	
1	E	103	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	103	

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
3	GOL	A	203	-	-	X	-
3	GOL	E	201	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 4986 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 MmpS5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	92	699	438	120	139	2	0	2	0
1	B	92	702	440	119	141	2	0	4	0
1	C	92	698	439	119	138	2	0	2	0
1	D	92	697	438	119	138	2	0	2	0
1	E	92	715	449	119	145	2	0	7	0
1	F	92	703	441	120	140	2	0	3	0

There are 30 discrepancies between the modelled and reference sequences:

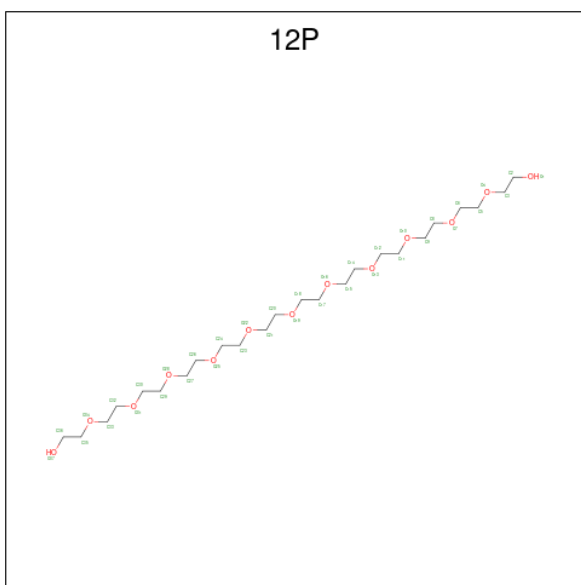
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP G7CDU2
A	2	PRO	-	expression tag	UNP G7CDU2
A	3	LEU	-	expression tag	UNP G7CDU2
A	4	GLY	-	expression tag	UNP G7CDU2
A	5	SER	-	expression tag	UNP G7CDU2
B	1	GLY	-	expression tag	UNP G7CDU2
B	2	PRO	-	expression tag	UNP G7CDU2
B	3	LEU	-	expression tag	UNP G7CDU2
B	4	GLY	-	expression tag	UNP G7CDU2
B	5	SER	-	expression tag	UNP G7CDU2
C	1	GLY	-	expression tag	UNP G7CDU2
C	2	PRO	-	expression tag	UNP G7CDU2
C	3	LEU	-	expression tag	UNP G7CDU2
C	4	GLY	-	expression tag	UNP G7CDU2
C	5	SER	-	expression tag	UNP G7CDU2
D	41	GLY	-	expression tag	UNP G7CDU2
D	42	PRO	-	expression tag	UNP G7CDU2

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	43	LEU	-	expression tag	UNP G7CDU2
D	44	GLY	-	expression tag	UNP G7CDU2
D	45	SER	-	expression tag	UNP G7CDU2
E	1	GLY	-	expression tag	UNP G7CDU2
E	2	PRO	-	expression tag	UNP G7CDU2
E	3	LEU	-	expression tag	UNP G7CDU2
E	4	GLY	-	expression tag	UNP G7CDU2
E	5	SER	-	expression tag	UNP G7CDU2
F	1	GLY	-	expression tag	UNP G7CDU2
F	2	PRO	-	expression tag	UNP G7CDU2
F	3	LEU	-	expression tag	UNP G7CDU2
F	4	GLY	-	expression tag	UNP G7CDU2
F	5	SER	-	expression tag	UNP G7CDU2

- Molecule 2 is DODECAETHYLENE GLYCOL (three-letter code: 12P) (formula:  $C_{24}H_{50}O_{13}$ ).



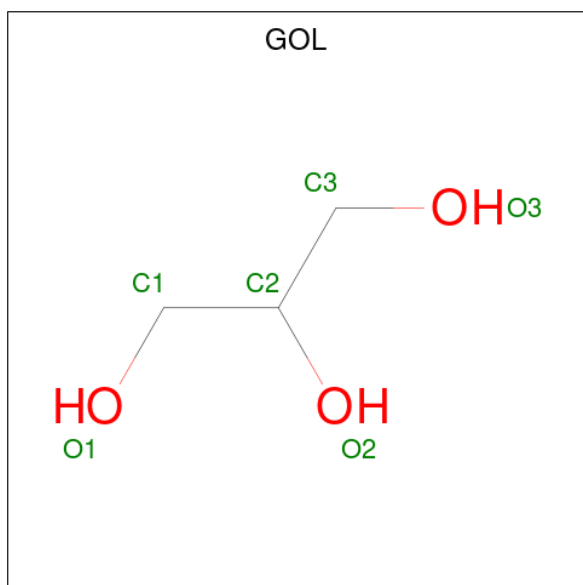
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			34	22	12		
2	B	1	Total	C	O	0	0
			19	12	7		
2	B	1	Total	C	O	0	0
			10	6	4		
2	C	1	Total	C	O	0	0
			7	4	3		
2	C	1	Total	C	O	0	0
			22	14	8		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	1	Total	C	O	0	0
			7	4	3		
2	E	1	Total	C	O	0	0
			7	4	3		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



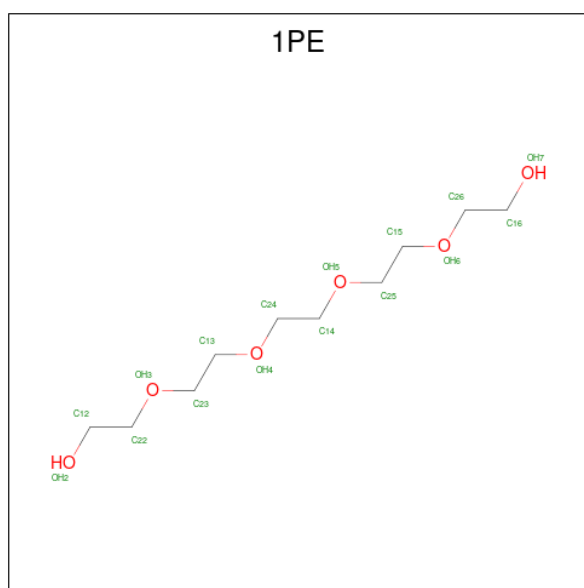
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total	C	O	0	0
			6	3	3		
3	E	1	Total	C	O	0	0
			6	3	3		
3	E	1	Total	C	O	0	0
			6	3	3		
3	E	1	Total	C	O	0	0
			6	3	3		
3	E	1	Total	C	O	0	0
			6	3	3		
3	F	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C<sub>10</sub>H<sub>22</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			16	10	6		

- Molecule 5 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	2	Total	I	0	2
			2	2		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	D	2	Total I 2 2	0	2
5	F	4	Total I 4 4	0	4

- Molecule 6 is water.


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	84	Total O 84 84	0	0
6	B	92	Total O 92 92	0	0
6	C	96	Total O 96 96	0	0
6	D	101	Total O 101 101	0	0
6	E	84	Total O 84 84	0	0
6	F	89	Total O 89 89	0	0

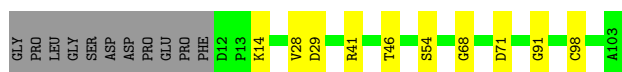


### 3 Residue-property plots [i](#)


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.

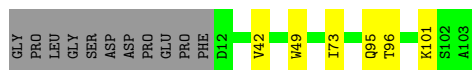
- Molecule 1: MmpS5

Chain A: 




- Molecule 1: MmpS5

Chain B: 




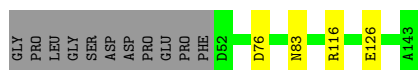
- Molecule 1: MmpS5

Chain C: 



- Molecule 1: MmpS5

Chain D: 




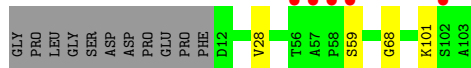
- Molecule 1: MmpS5

Chain E: 



- Molecule 1: MmpS5

Chain F:  5% 85% 11%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	50.28Å 85.26Å 89.12Å 90.00° 95.19° 90.00°	Depositor
Resolution (Å)	40.08 – 1.95 40.08 – 1.95	Depositor EDS
% Data completeness (in resolution range)	94.5 (40.08-1.95) 94.5 (40.08-1.95)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.64 (at 1.95Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.161 , 0.194 0.161 , 0.193	Depositor DCC
$R_{free}$ test set	5176 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.3	Xtrriage
Anisotropy	0.335	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 59.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4986	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.66% 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: IOD, GOL, 12P, 1PE

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	0.50	1/712 (0.1%)	0.72	0/974
1	B	0.46	0/721	0.70	0/986
1	C	0.42	0/711	0.73	0/973
1	D	0.46	0/710	0.70	0/972
1	E	0.43	0/743	0.71	0/1016
1	F	0.44	0/717	0.69	0/982
All	All	0.45	1/4314 (0.0%)	0.71	0/5903

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	46	THR	C-N	5.63	1.47	1.34

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	699	0	693	7	0
1	B	702	0	700	6	1
1	C	698	0	696	5	0
1	D	697	0	694	3	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	715	0	718	13	1
1	F	703	0	688	3	0
2	A	34	0	45	3	0
2	B	29	0	37	4	0
2	C	29	0	37	2	0
2	D	7	0	8	0	0
2	E	7	0	8	0	0
3	A	30	0	40	7	0
3	B	12	0	15	0	0
3	C	6	0	8	0	0
3	D	12	0	16	0	0
3	E	30	0	39	5	0
3	F	6	0	8	1	0
4	A	16	0	22	1	0
5	C	2	0	0	0	0
5	D	2	0	0	1	0
5	F	4	0	0	0	0
6	A	84	0	0	2	0
6	B	92	0	0	1	0
6	C	96	0	0	0	0
6	D	101	0	0	3	0
6	E	84	0	0	4	0
6	F	89	0	0	1	0
All	All	4986	0	4472	44	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All (44) 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:B:95:GLN:NE2	6:B:301:HOH:O	2.09	0.85
1:A:41:ARG:HH22	3:A:203:GOL:H2	1.47	0.78
1:E:71:ASP:OD1	6:E:301:HOH:O	2.02	0.77
1:D:76:ASP:OD2	6:D:301:HOH:O	2.04	0.75
1:D:83:ASN:OD1	6:D:302:HOH:O	2.09	0.69
1:E:95:GLN:NE2	6:E:303:HOH:O	2.22	0.67
5:D:204[L]:IOD:I	6:D:375:HOH:O	2.86	0.61
1:B:101:LYS:HE2	2:B:202:12P:H52	1.82	0.61
3:A:202:GOL:O2	6:A:301:HOH:O	2.17	0.59
1:F:101:LYS:NZ	6:F:301:HOH:O	2.34	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:28:VAL:HG12	1:F:68:GLY:HA3	1.86	0.58
3:A:202:GOL:H32	1:C:95:GLN:O	2.07	0.55
2:A:201:12P:H81	1:C:97:TYR:CZ	2.43	0.54
1:F:101:LYS:HE2	3:F:201:GOL:O2	2.09	0.53
1:A:98:CYS:HA	3:A:202:GOL:H12	1.90	0.53
1:B:49:TRP:HE1	2:B:204:12P:H201	1.71	0.52
3:A:203:GOL:H31	6:A:308:HOH:O	2.09	0.51
1:E:49:TRP:CE2	3:E:205:GOL:H12	2.47	0.50
1:B:73:ILE:HG23	1:B:96:THR:HB	1.93	0.50
1:E:38:ARG:NH1	6:E:304:HOH:O	2.29	0.49
1:E:49:TRP:NE1	3:E:205:GOL:H12	2.27	0.48
1:B:42:VAL:HG11	2:B:204:12P:H212	1.96	0.47
1:E:51:ILE:HD12	3:E:205:GOL:H11	1.98	0.46
1:C:38:ARG:HE	1:C:38:ARG:HB2	1.50	0.45
1:E:27:VAL:HG22	1:E:46[B]:THR:HG22	1.99	0.45
1:E:84:ARG:NH1	1:E:100[B]:GLU:OE2	2.41	0.45
1:E:101:LYS:NZ	6:E:302:HOH:O	2.21	0.45
1:A:28:VAL:HA	1:A:68:GLY:HA3	1.99	0.45
2:A:201:12P:H151	2:A:201:12P:H122	1.31	0.44
1:B:49:TRP:NE1	2:B:204:12P:H201	2.33	0.43
2:A:201:12P:H201	2:A:201:12P:H172	1.55	0.43
1:C:99[A]:ILE:CG2	2:C:403:12P:H232	2.48	0.42
1:E:36:ASP:OD1	3:E:201:GOL:H12	2.20	0.42
1:E:41:ARG:HE	1:E:43:ASN:HD21	1.65	0.42
2:C:402:12P:O25	2:C:403:12P:H362	2.20	0.41
1:E:31:ASN:HD22	3:E:202:GOL:H12	1.85	0.41
4:A:205:1PE:H231	4:A:205:1PE:H122	1.84	0.41
1:A:41:ARG:NH2	3:A:203:GOL:H2	2.24	0.41
1:C:77:ILE:HB	1:C:85:VAL:HG22	2.02	0.41
1:A:71:ASP:HA	1:A:91:GLY:O	2.21	0.41
1:A:14:LYS:O	1:A:54:SER:HA	2.20	0.40
1:E:12:ASP:HA	1:E:13:PRO:HD3	1.97	0.40
1:D:116:ARG:HD3	1:D:126:GLU:OE2	2.20	0.40
1:A:29:ASP:HB2	3:A:203:GOL:H11	2.03	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:95:GLN:OE1	1:E:97:TYR:OH[2_647]	2.12	0.08

## 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	91/103 (88%)	88 (97%)	3 (3%)	0	100	100
1	B	93/103 (90%)	90 (97%)	3 (3%)	0	100	100
1	C	91/103 (88%)	88 (97%)	3 (3%)	0	100	100
1	D	91/103 (88%)	89 (98%)	2 (2%)	0	100	100
1	E	96/103 (93%)	94 (98%)	2 (2%)	0	100	100
1	F	92/103 (89%)	90 (98%)	2 (2%)	0	100	100
All	All	554/618 (90%)	539 (97%)	15 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	77/86 (90%)	77 (100%)	0	100	100
1	B	79/86 (92%)	79 (100%)	0	100	100
1	C	77/86 (90%)	77 (100%)	0	100	100
1	D	77/86 (90%)	77 (100%)	0	100	100
1	E	82/86 (95%)	81 (99%)	1 (1%)	71	68
1	F	77/86 (90%)	76 (99%)	1 (1%)	69	65
All	All	469/516 (91%)	467 (100%)	2 (0%)	91	90

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	45	VAL
1	F	59	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	31	ASN
1	E	43	ASN
1	F	43	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 32 ligands modelled in this entry, 8 are monoatomic - leaving 24 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	GOL	E	203	-	5,5,5	1.17	1 (20%)	5,5,5	0.93	0
3	GOL	C	401	-	5,5,5	1.01	0	5,5,5	0.93	0
3	GOL	E	201	-	5,5,5	0.73	0	5,5,5	1.10	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	12P	D	202	-	6,6,36	0.14	0	5,5,35	0.22	0
3	GOL	B	203	-	5,5,5	0.96	0	5,5,5	0.85	0
3	GOL	E	204	-	5,5,5	0.96	0	5,5,5	0.90	0
3	GOL	A	204	-	5,5,5	0.80	0	5,5,5	0.97	0
3	GOL	E	205	-	5,5,5	0.90	0	5,5,5	0.95	0
3	GOL	A	203	-	5,5,5	0.69	0	5,5,5	0.95	0
2	12P	A	201	-	33,33,36	0.56	0	32,32,35	0.48	0
2	12P	E	206	-	6,6,36	0.18	0	5,5,35	0.21	0
3	GOL	F	201	-	5,5,5	0.97	0	5,5,5	0.94	0
3	GOL	A	202	-	5,5,5	0.95	0	5,5,5	1.22	1 (20%)
3	GOL	D	201	-	5,5,5	1.00	0	5,5,5	0.90	0
3	GOL	A	207	-	5,5,5	0.94	0	5,5,5	0.87	0
3	GOL	A	206	-	5,5,5	0.99	0	5,5,5	1.21	0
2	12P	B	204	-	9,9,36	0.29	0	8,8,35	0.45	0
2	12P	C	403	-	21,21,36	0.10	0	20,20,35	0.25	0
3	GOL	E	202	-	5,5,5	1.19	1 (20%)	5,5,5	1.12	1 (20%)
2	12P	B	202	-	18,18,36	0.31	0	17,17,35	0.28	0
3	GOL	D	203	-	5,5,5	1.11	1 (20%)	5,5,5	0.84	0
3	GOL	B	201	-	5,5,5	1.27	1 (20%)	5,5,5	0.81	0
4	1PE	A	205	-	15,15,15	0.18	0	14,14,14	0.17	0
2	12P	C	402	-	6,6,36	0.12	0	5,5,35	0.17	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	E	203	-	-	4/4/4/4	-
3	GOL	C	401	-	-	0/4/4/4	-
3	GOL	E	201	-	-	2/4/4/4	-
2	12P	D	202	-	-	2/4/4/34	-
3	GOL	B	203	-	-	2/4/4/4	-
3	GOL	E	204	-	-	1/4/4/4	-
3	GOL	A	204	-	-	4/4/4/4	-
3	GOL	E	205	-	-	4/4/4/4	-
3	GOL	A	203	-	-	0/4/4/4	-
2	12P	A	201	-	-	17/31/31/34	-
2	12P	E	206	-	-	2/4/4/34	-
3	GOL	F	201	-	-	4/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	202	-	-	4/4/4/4	-
3	GOL	D	201	-	-	2/4/4/4	-
3	GOL	A	207	-	-	2/4/4/4	-
3	GOL	A	206	-	-	1/4/4/4	-
2	12P	B	204	-	-	3/7/7/34	-
2	12P	C	403	-	-	8/19/19/34	-
3	GOL	E	202	-	-	2/4/4/4	-
2	12P	B	202	-	-	6/16/16/34	-
3	GOL	D	203	-	-	2/4/4/4	-
3	GOL	B	201	-	-	2/4/4/4	-
4	1PE	A	205	-	-	6/13/13/13	-
2	12P	C	402	-	-	1/4/4/34	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	201	GOL	O2-C2	-2.42	1.36	1.43
3	E	202	GOL	O2-C2	-2.29	1.36	1.43
3	E	203	GOL	O2-C2	-2.13	1.37	1.43
3	D	203	GOL	O2-C2	-2.01	1.37	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	A	202	GOL	C3-C2-C1	-2.31	102.73	111.70
3	E	202	GOL	C3-C2-C1	-2.12	103.47	111.70

There are no chirality outliers.

All (81) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	202	GOL	O1-C1-C2-C3
3	A	204	GOL	O1-C1-C2-C3
3	D	201	GOL	O1-C1-C2-C3
3	D	203	GOL	C1-C2-C3-O3
3	E	201	GOL	O1-C1-C2-C3
3	E	202	GOL	O1-C1-C2-C3
3	E	205	GOL	O1-C1-C2-C3
3	F	201	GOL	C1-C2-C3-O3

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	A	201	12P	C15-C14-O13-C12
2	A	201	12P	C17-C18-O19-C20
2	A	201	12P	C18-C17-O16-C15
4	A	205	1PE	OH6-C15-C25-OH5
2	A	201	12P	O10-C11-C12-O13
2	E	206	12P	O22-C23-C24-O25
2	A	201	12P	O19-C20-C21-O22
2	A	201	12P	O25-C26-C27-O28
2	A	201	12P	O7-C8-C9-O10
2	A	201	12P	O13-C14-C15-O16
3	A	202	GOL	O2-C2-C3-O3
3	B	203	GOL	O1-C1-C2-O2
3	E	205	GOL	O1-C1-C2-O2
3	E	205	GOL	O2-C2-C3-O3
4	A	205	1PE	OH7-C16-C26-OH6
3	A	202	GOL	C1-C2-C3-O3
3	A	204	GOL	C1-C2-C3-O3
3	A	206	GOL	C1-C2-C3-O3
3	A	207	GOL	O1-C1-C2-C3
3	B	203	GOL	O1-C1-C2-C3
3	E	203	GOL	O1-C1-C2-C3
3	E	205	GOL	C1-C2-C3-O3
3	F	201	GOL	O1-C1-C2-C3
3	A	204	GOL	O1-C1-C2-O2
3	D	201	GOL	O1-C1-C2-O2
3	D	203	GOL	O2-C2-C3-O3
3	E	201	GOL	O1-C1-C2-O2
3	E	202	GOL	O1-C1-C2-O2
3	F	201	GOL	O2-C2-C3-O3
2	A	201	12P	O4-C5-C6-O7
2	C	403	12P	O34-C35-C36-O37
2	B	202	12P	O4-C5-C6-O7
2	C	402	12P	O25-C26-C27-O28
2	D	202	12P	O22-C23-C24-O25
3	E	203	GOL	O1-C1-C2-O2
4	A	205	1PE	OH2-C12-C22-OH3
3	F	201	GOL	O1-C1-C2-O2
2	B	202	12P	C2-C3-O4-C5
2	B	204	12P	C20-C21-O22-C23
2	B	204	12P	C24-C23-O22-C21
2	C	403	12P	C17-C18-O19-C20
2	A	201	12P	C11-C12-O13-C14

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	D	202	12P	C27-C26-O25-C24
2	A	201	12P	C24-C23-O22-C21
2	A	201	12P	C21-C20-O19-C18
2	C	403	12P	C27-C26-O25-C24
3	A	204	GOL	O2-C2-C3-O3
2	C	403	12P	O22-C23-C24-O25
4	A	205	1PE	C24-C14-OH5-C25
2	C	403	12P	C21-C20-O19-C18
4	A	205	1PE	C12-C22-OH3-C23
2	A	201	12P	C12-C11-O10-C9
2	C	403	12P	C32-C33-O34-C35
2	B	202	12P	C18-C17-O16-C15
2	B	202	12P	C6-C5-O4-C3
3	B	201	GOL	O1-C1-C2-O2
2	A	201	12P	C27-C26-O25-C24
4	A	205	1PE	C25-C15-OH6-C26
2	B	202	12P	C5-C6-O7-C8
3	A	202	GOL	O1-C1-C2-O2
3	A	207	GOL	O1-C1-C2-O2
2	A	201	12P	C33-C32-O31-C30
2	A	201	12P	C5-C6-O7-C8
2	A	201	12P	O16-C17-C18-O19
3	E	203	GOL	C1-C2-C3-O3
2	C	403	12P	C20-C21-O22-C23
3	E	203	GOL	O2-C2-C3-O3
3	E	204	GOL	O2-C2-C3-O3
2	E	206	12P	C20-C21-O22-C23
3	B	201	GOL	O1-C1-C2-C3
2	B	202	12P	O16-C17-C18-O19
2	C	403	12P	O16-C17-C18-O19
2	B	204	12P	O22-C23-C24-O25

There are no ring outliers.

12 monomers are involved in 23 short contacts:

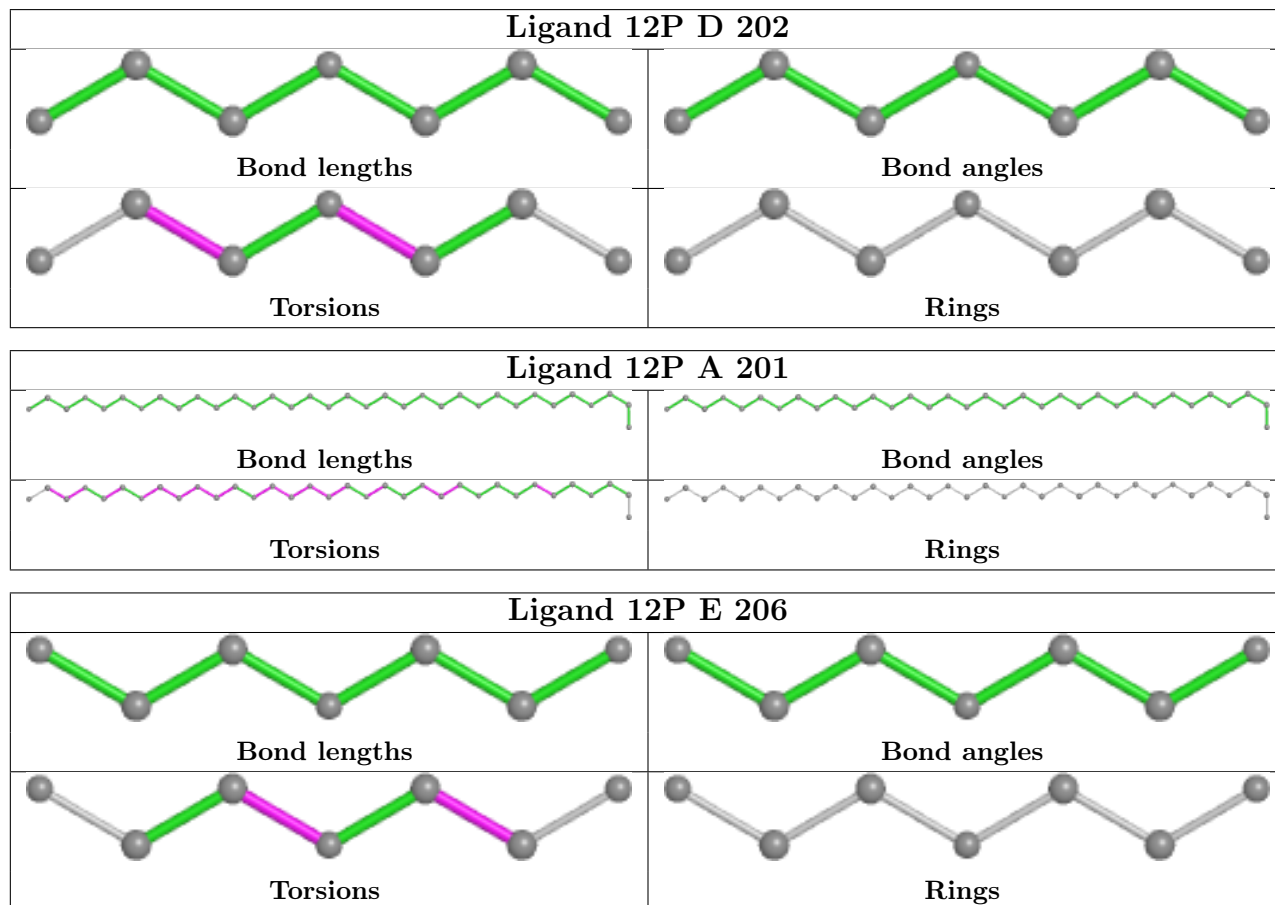
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	201	GOL	1	0
3	E	205	GOL	3	0
3	A	203	GOL	4	0
2	A	201	12P	3	0
3	F	201	GOL	1	0
3	A	202	GOL	3	0

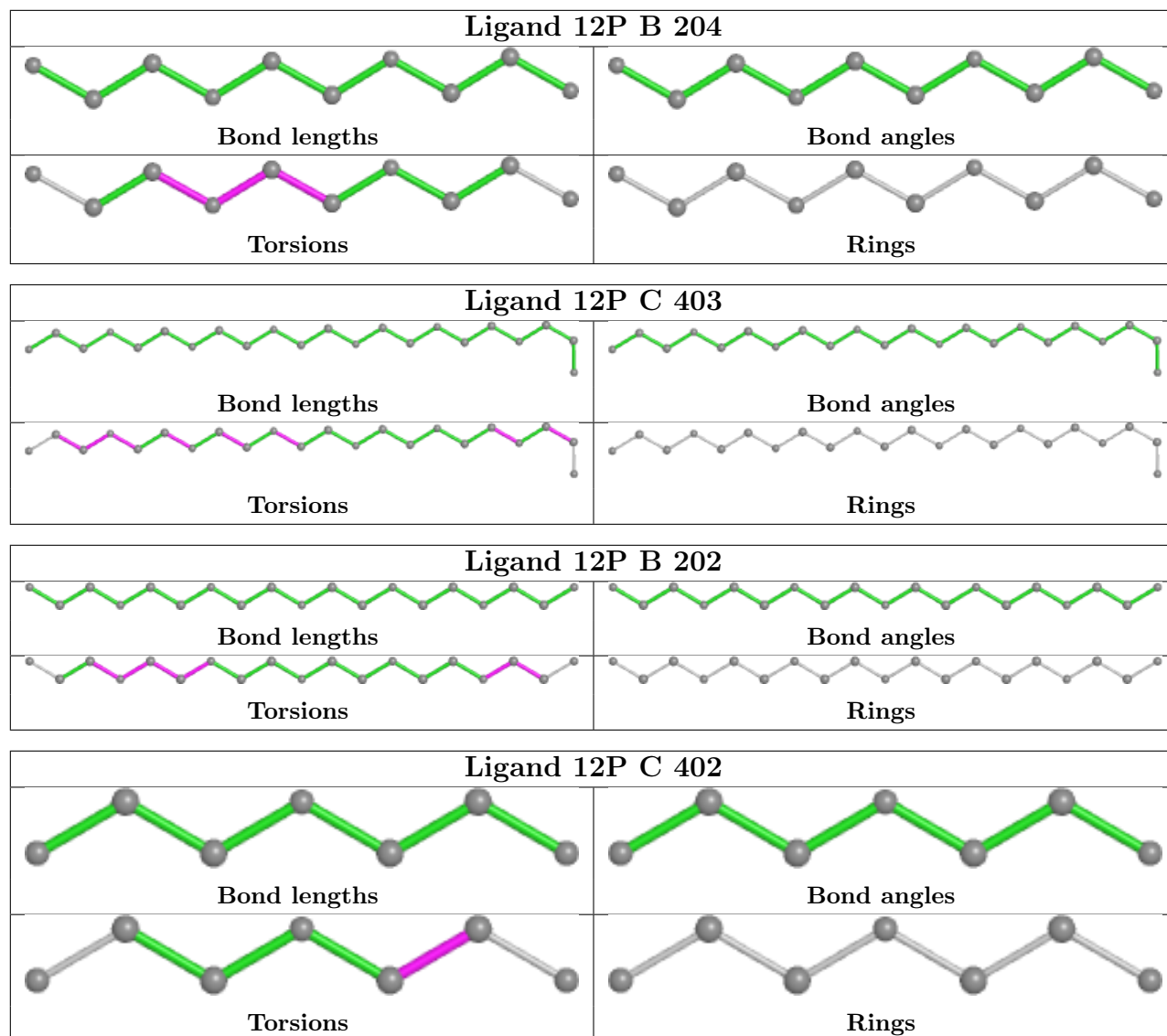
*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	204	12P	3	0
2	C	403	12P	2	0
3	E	202	GOL	1	0
2	B	202	12P	1	0
4	A	205	1PE	1	0
2	C	402	12P	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 [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	92/103 (89%)	-0.29	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	12, 18, 28, 40	0
1	B	92/103 (89%)	-0.21	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	11, 17, 28, 39	0
1	C	92/103 (89%)	0.00	1 (1%) <span style="border: 1px solid blue; padding: 2px;">80</span> <span style="border: 1px solid blue; padding: 2px;">85</span>	13, 19, 29, 46	0
1	D	92/103 (89%)	-0.18	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	12, 19, 33, 44	0
1	E	92/103 (89%)	0.01	2 (2%) <span style="border: 1px solid blue; padding: 2px;">62</span> <span style="border: 1px solid blue; padding: 2px;">70</span>	15, 22, 36, 48	0
1	F	92/103 (89%)	0.03	5 (5%) <span style="border: 1px solid red; padding: 2px;">25</span> <span style="border: 1px solid red; padding: 2px;">34</span>	14, 22, 35, 46	0
All	All	552/618 (89%)	-0.11	8 (1%) <span style="border: 1px solid blue; padding: 2px;">75</span> <span style="border: 1px solid blue; padding: 2px;">82</span>	11, 19, 33, 48	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	56	THR	3.5
1	F	57	ALA	3.4
1	E	58	PRO	2.5
1	F	58	PRO	2.2
1	E	56	THR	2.2
1	F	59	SER	2.2
1	C	99[A]	ILE	2.1
1	F	102	SER	2.1

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

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

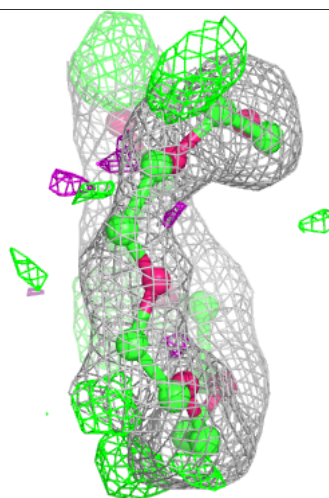
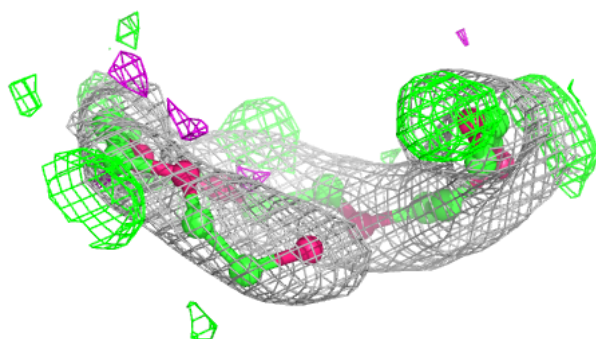
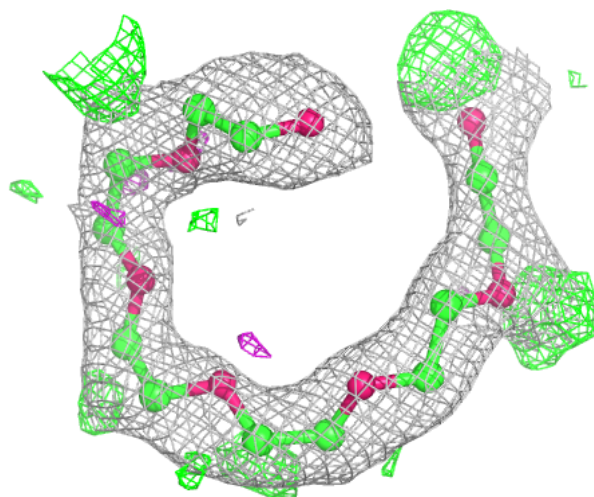
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	12P	B	202	19/37	0.75	0.21	27,42,51,52	0
2	12P	B	204	10/37	0.75	0.33	29,44,54,61	0
3	GOL	D	201	6/6	0.75	0.24	51,54,60,70	0
4	1PE	A	205	16/16	0.79	0.31	38,51,58,64	0
3	GOL	E	205	6/6	0.80	0.28	39,47,49,60	0
3	GOL	E	201	6/6	0.80	0.51	50,53,63,64	0
3	GOL	A	207	6/6	0.81	0.18	40,48,54,55	0
3	GOL	C	401	6/6	0.81	0.25	41,51,56,57	0
2	12P	D	202	7/37	0.82	0.18	44,48,51,59	0
2	12P	E	206	7/37	0.82	0.23	41,48,53,61	0
3	GOL	F	201	6/6	0.83	0.26	50,65,67,70	0
3	GOL	B	201	6/6	0.83	0.20	37,40,45,46	0
3	GOL	A	203	6/6	0.84	0.15	39,44,44,60	0
3	GOL	B	203	6/6	0.85	0.21	35,47,53,64	0
3	GOL	E	204	6/6	0.86	0.28	34,39,45,53	6
2	12P	A	201	34/37	0.87	0.22	26,49,67,69	0
3	GOL	E	202	6/6	0.87	0.17	36,42,52,62	0
3	GOL	D	203	6/6	0.89	0.18	31,46,51,59	0
3	GOL	A	206	6/6	0.90	0.13	23,37,41,42	0
2	12P	C	402	7/37	0.92	0.17	23,37,44,54	0
3	GOL	A	204	6/6	0.92	0.19	27,40,45,51	0
3	GOL	A	202	6/6	0.92	0.13	23,30,38,40	0
3	GOL	E	203	6/6	0.93	0.23	28,31,39,41	0
2	12P	C	403	22/37	0.94	0.11	26,42,55,56	0
5	IOD	C	404[L]	1/1	0.96	0.13	45,45,45,45	1
5	IOD	D	204[L]	1/1	0.97	0.05	52,52,52,52	1
5	IOD	C	405[L]	1/1	0.99	0.09	24,24,24,24	0
5	IOD	F	202[L]	1/1	0.99	0.09	23,23,23,23	0
5	IOD	F	203[L]	1/1	0.99	0.08	21,21,21,21	1
5	IOD	F	204[L]	1/1	0.99	0.07	20,20,20,20	1
5	IOD	F	205[L]	1/1	0.99	0.09	21,21,21,21	1
5	IOD	D	205[L]	1/1	1.00	0.08	20,20,20,20	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.



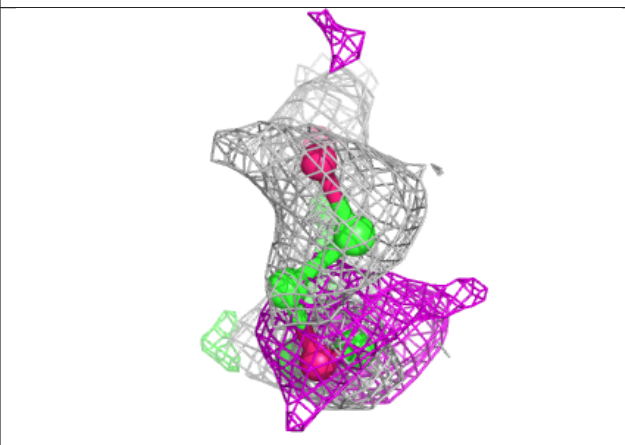
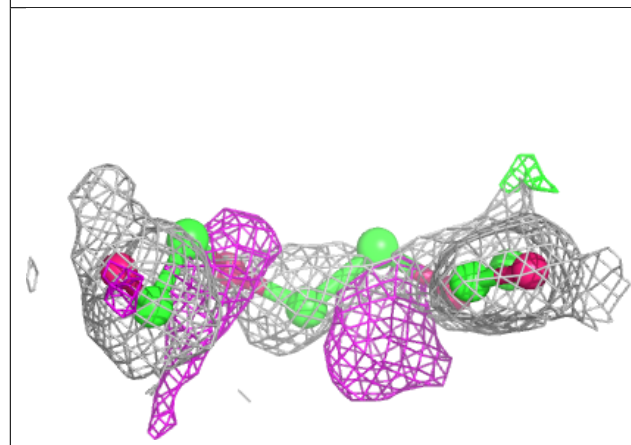
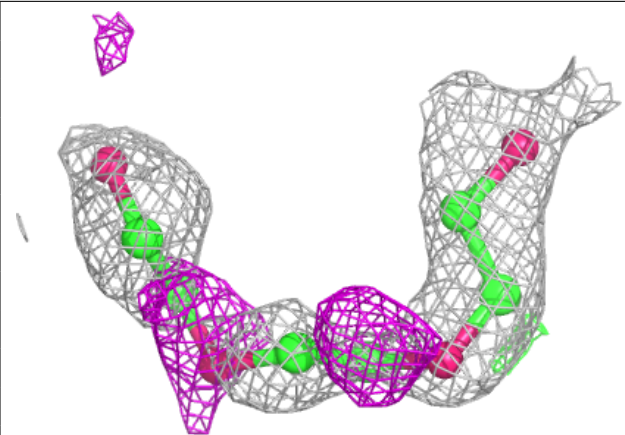
**Electron density around 12P B 202:**

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

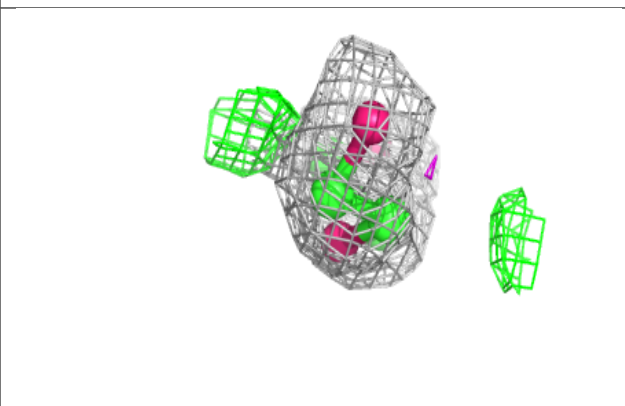
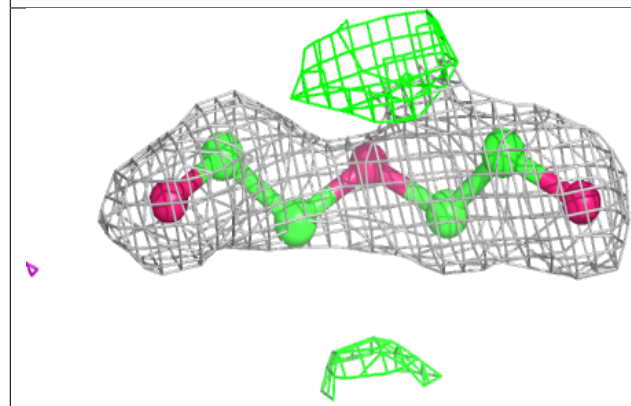
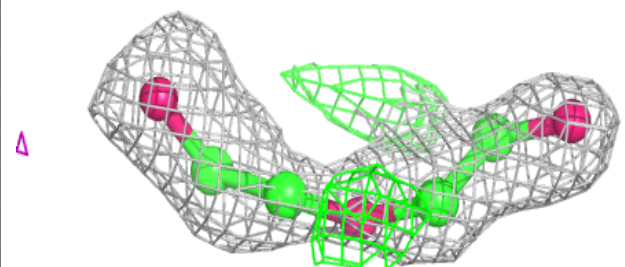


**Electron density around 12P B 204:**

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

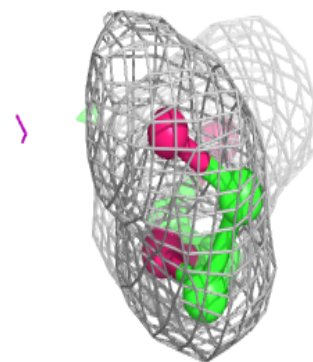
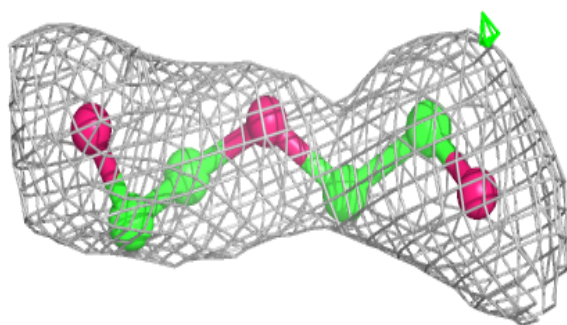
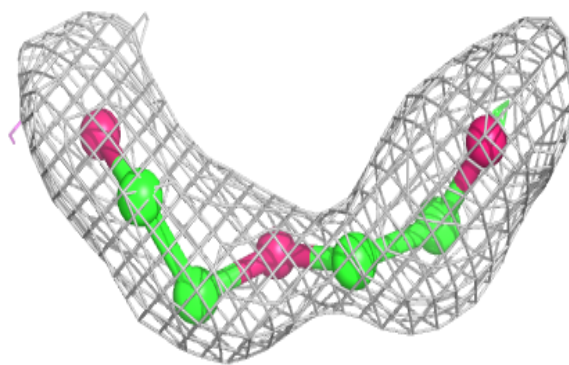
**Electron density around 12P D 202:**

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

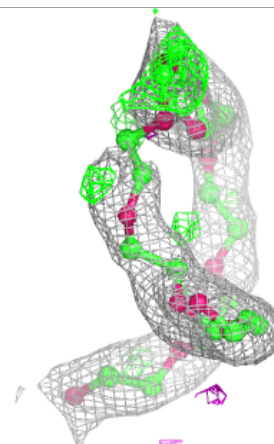
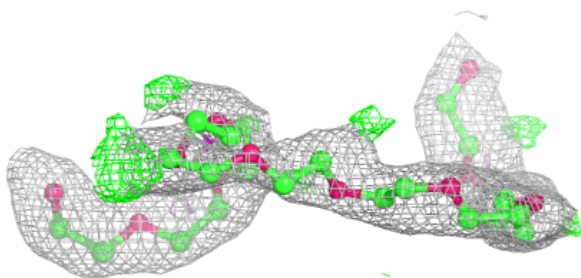
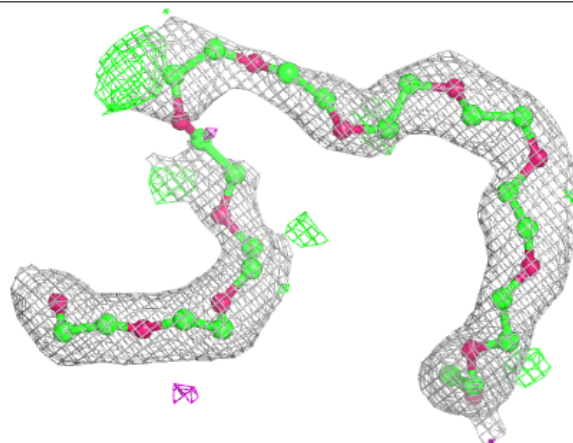


**Electron density around 12P E 206:**

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

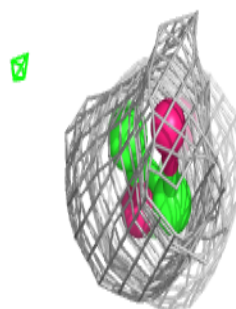
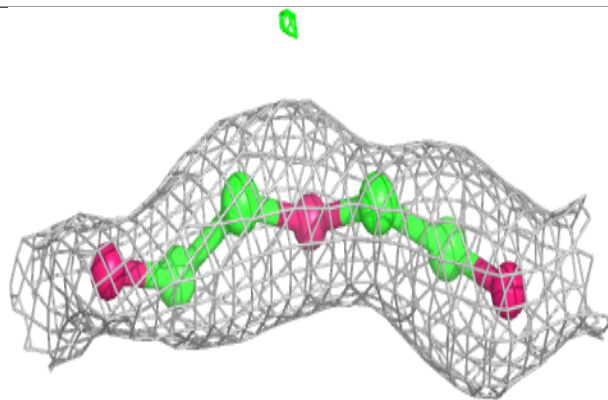
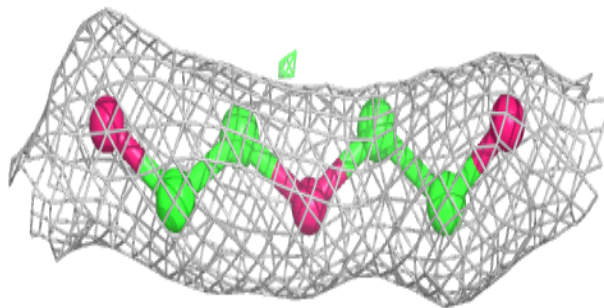
**Electron density around 12P A 201:**

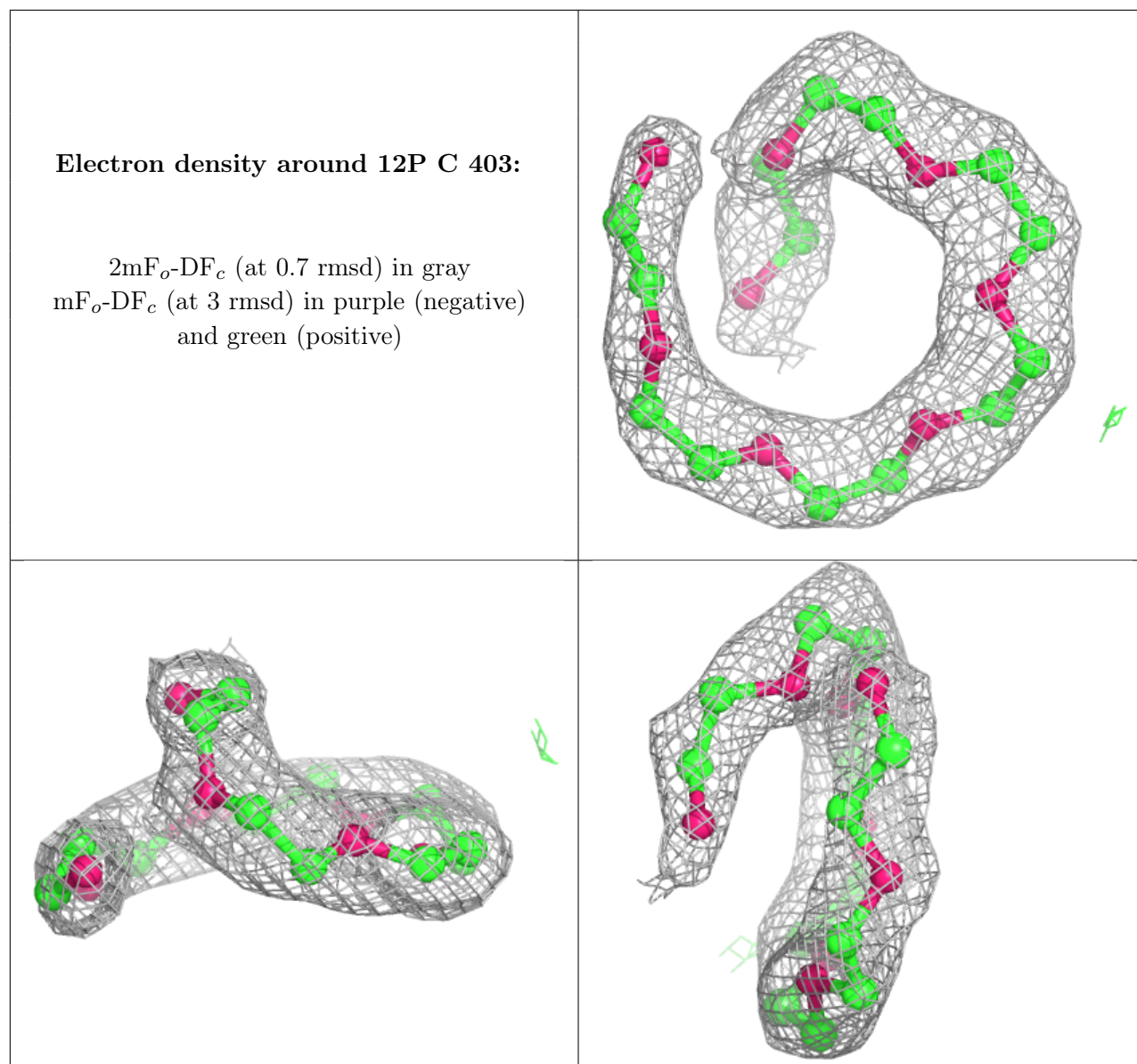
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around 12P C 402:**

$2mF_o-DF_c$  (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.