



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 10, 2020 – 12:45 PM BST

PDB ID : 5HPO
Title : Cycloalternan-forming enzyme from *Listeria monocytogenes* in complex with maltopentaose
Authors : Halavaty, A.S.; Light, S.H.; Minasov, G.; Winsor, J.; Grimshaw, S.; Shuvalova, L.; Peterson, S.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on : 2016-01-20
Resolution : 1.80 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

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	:	2.13.1
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)	:	Engl & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

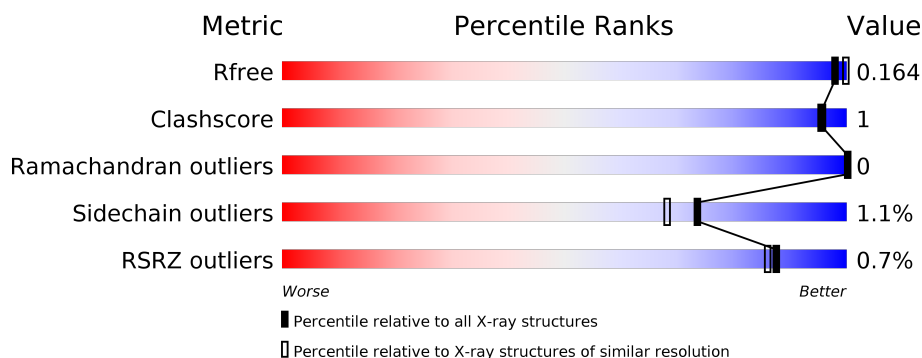
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 ≥ 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	1084	<div> <div style="width: 95%;"></div> <div>95%</div> </div>
2	B	4	<div> <div style="width: 25%; background-color: green;"></div> <div style="width: 75%; background-color: yellow;"></div> <div>25%</div> <div>75%</div> </div>
2	C	4	<div> <div style="width: 25%; background-color: green;"></div> <div style="width: 75%; background-color: yellow;"></div> <div>25%</div> <div>75%</div> </div>
3	D	3	<div> <div style="width: 33%; background-color: green;"></div> <div style="width: 67%; background-color: yellow;"></div> <div>33%</div> <div>67%</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
7	PEG	A	1106	-	-	X	-

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 9939 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 Lmo2446 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	1059	8560	5363	1365	1799	33	0	30	0

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-23	MSE	-	initiating methionine	UNP Q8Y4J2
A	-22	HIS	-	expression tag	UNP Q8Y4J2
A	-21	HIS	-	expression tag	UNP Q8Y4J2
A	-20	HIS	-	expression tag	UNP Q8Y4J2
A	-19	HIS	-	expression tag	UNP Q8Y4J2
A	-18	HIS	-	expression tag	UNP Q8Y4J2
A	-17	HIS	-	expression tag	UNP Q8Y4J2
A	-16	SER	-	expression tag	UNP Q8Y4J2
A	-15	SER	-	expression tag	UNP Q8Y4J2
A	-14	GLY	-	expression tag	UNP Q8Y4J2
A	-13	VAL	-	expression tag	UNP Q8Y4J2
A	-12	ASP	-	expression tag	UNP Q8Y4J2
A	-11	LEU	-	expression tag	UNP Q8Y4J2
A	-10	GLY	-	expression tag	UNP Q8Y4J2
A	-9	THR	-	expression tag	UNP Q8Y4J2
A	-8	GLU	-	expression tag	UNP Q8Y4J2
A	-7	ASN	-	expression tag	UNP Q8Y4J2
A	-6	LEU	-	expression tag	UNP Q8Y4J2
A	-5	TYR	-	expression tag	UNP Q8Y4J2
A	-4	PHE	-	expression tag	UNP Q8Y4J2
A	-3	GLN	-	expression tag	UNP Q8Y4J2
A	-2	SER	-	expression tag	UNP Q8Y4J2
A	-1	ASN	-	expression tag	UNP Q8Y4J2

- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	B	4	Total	C	O	0	0	1
			34	18	16			
2	C	4	Total	C	O	0	0	1
			34	18	16			

- Molecule 3 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
3	D	3	Total	C	O	0	0	1
			23	12	11			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Mg	0	0
			2	2		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Ca	0	0
			1	1		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	Cl	0	0
			2	2		

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



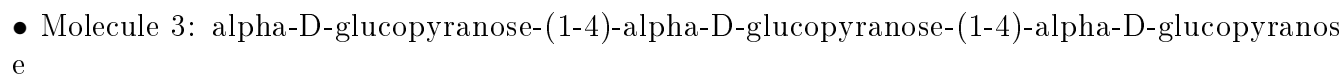
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	1259	Total	O	0	24
			1276	1276		



- Molecule 1: Lmo2446 protein



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	162.65Å 100.13Å 73.28Å 90.00° 105.47° 90.00°	Depositor
Resolution (Å)	29.92 – 1.80 29.92 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.3 (29.92-1.80) 99.4 (29.92-1.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.11 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.133 , 0.164 0.133 , 0.164	Depositor DCC
R_{free} test set	5194 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	20.5	Xtriage
Anisotropy	0.203	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 48.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	9939	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.24% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PEG, MG, GLC, CA, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.51	0/8740	0.77	4/11854 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	158	MSE	CG-SE-CE	-6.95	83.61	98.90
1	A	607	PHE	CB-CG-CD1	5.81	124.86	120.80
1	A	255	ASP	CB-CG-OD1	5.69	123.42	118.30
1	A	607	PHE	CB-CG-CD2	-5.21	117.16	120.80

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	8560	0	7836	19	0
2	B	34	0	28	0	0
2	C	34	0	28	0	0
3	D	23	0	19	0	0
4	A	2	0	0	0	0
5	A	1	0	0	0	0
6	A	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	7	0	10	4	0
8	A	1276	0	0	5	1
All	All	9939	0	7921	19	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 1.

The worst 5 of 19 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:846[B]:MSE:SE	8:A:2124:HOH:O	2.58	0.71
1:A:403:GLU:OE1	1:A:468:ARG:NH2	2.32	0.63
1:A:272:TYR:HB3	7:A:1106:PEG:H22	1.81	0.62
1:A:272:TYR:H	7:A:1106:PEG:C2	2.17	0.58
1:A:989:GLU:HG2	1:A:1027:GLU:OE1	2.05	0.56

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)
8:A:1235:HOH:O	8:A:1235:HOH:O[2_454]	1.97	0.23

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	1087/1084 (100%)	1058 (97%)	29 (3%)	0	100 100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	920/881 (104%)	910 (99%)	10 (1%)	73	68

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	596	TRP
1	A	632	PHE
1	A	778	GLU
1	A	533	THR
1	A	656	GLN

Some 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](#)

Of 11 monosaccharides modelled in this entry, 8 were used for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	GLC	B	2	2	11,11,12	0.35	0	15,15,17	1.20	2 (13%)
2	GLC	B	3	2	11,11,12	0.57	0	15,15,17	0.97	1 (6%)
2	GLC	B	4	2	11,11,12	0.57	0	15,15,17	1.27	2 (13%)
2	GLC	C	2	2	11,11,12	0.59	0	15,15,17	1.22	1 (6%)
2	GLC	C	3	2	11,11,12	0.37	0	15,15,17	0.98	1 (6%)
2	GLC	C	4	2	11,11,12	0.54	0	15,15,17	0.91	1 (6%)
3	GLC	D	2	3	11,11,12	0.76	0	15,15,17	1.32	2 (13%)
3	GLC	D	3	3	11,11,12	0.24	0	15,15,17	0.98	1 (6%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GLC	B	2	2	-	2/2/19/22	0/1/1/1
2	GLC	B	3	2	-	0/2/19/22	0/1/1/1
2	GLC	B	4	2	-	0/2/19/22	0/1/1/1
2	GLC	C	2	2	-	2/2/19/22	0/1/1/1
2	GLC	C	3	2	-	0/2/19/22	0/1/1/1
2	GLC	C	4	2	-	0/2/19/22	0/1/1/1
3	GLC	D	2	3	-	1/2/19/22	0/1/1/1
3	GLC	D	3	3	-	2/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2	GLC	C1-O5-C5	3.32	116.69	112.19
2	C	2	GLC	C1-O5-C5	3.07	116.36	112.19
2	B	4	GLC	C1-C2-C3	2.99	113.34	109.67
3	D	3	GLC	C1-O5-C5	2.99	116.24	112.19
3	D	2	GLC	C1-O5-C5	2.92	116.15	112.19

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	2	GLC	C4-C5-C6-O6

Continued on next page...

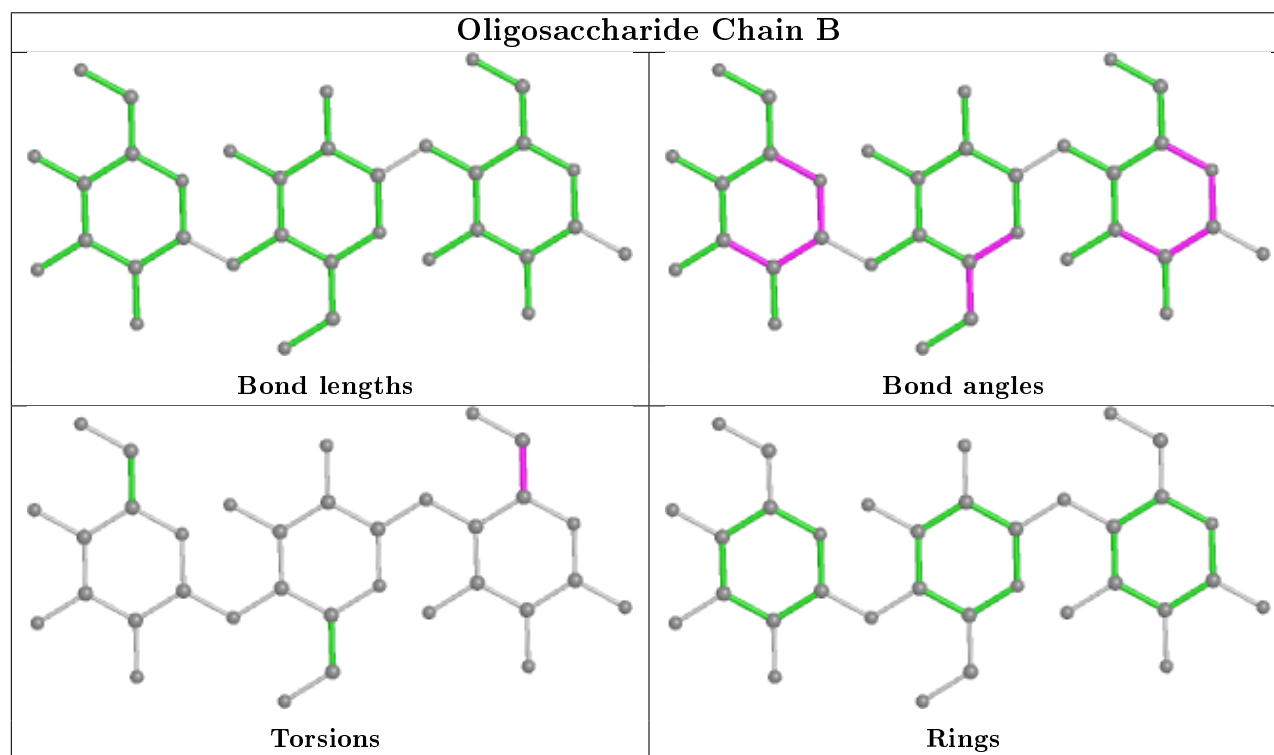
Continued from previous page...

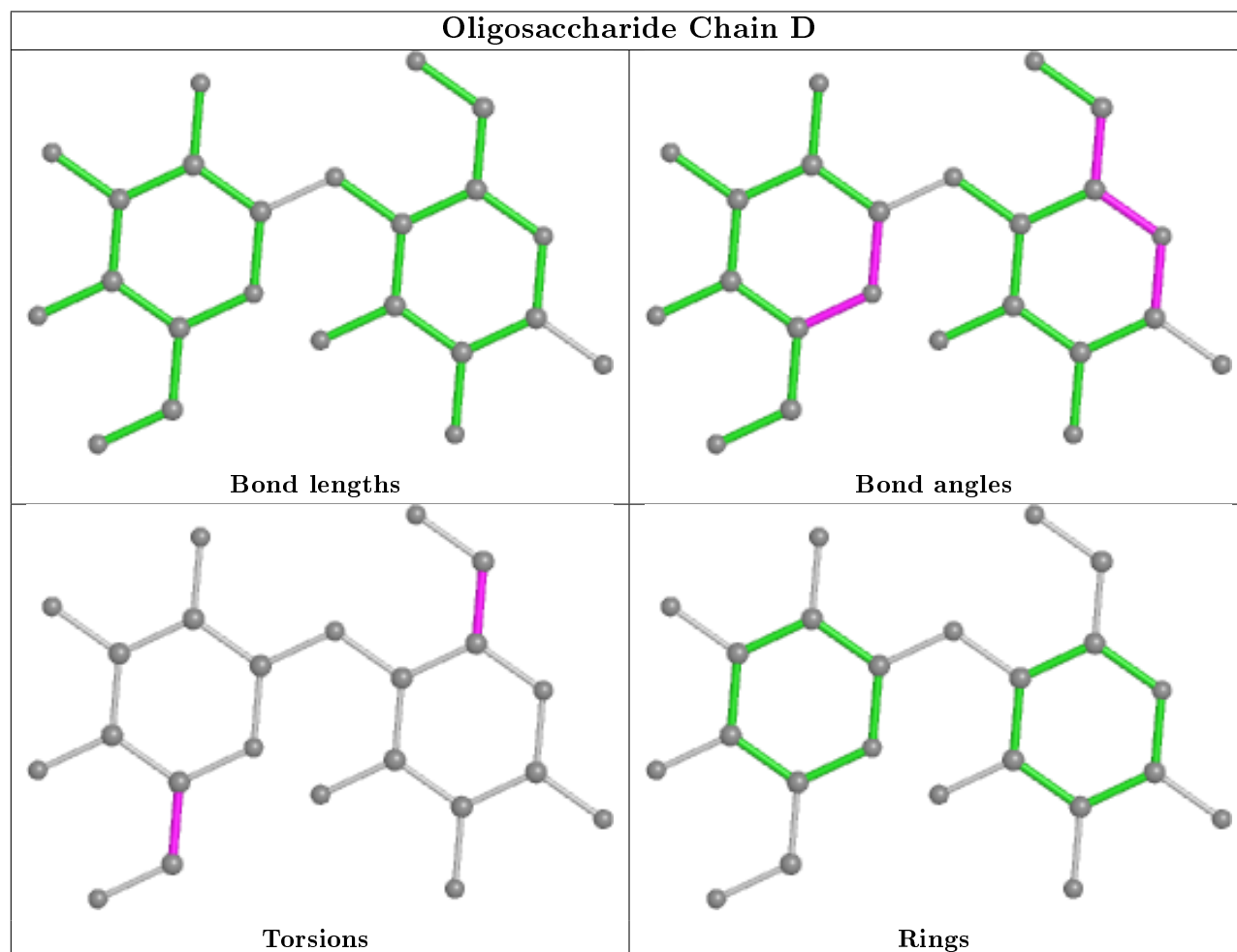
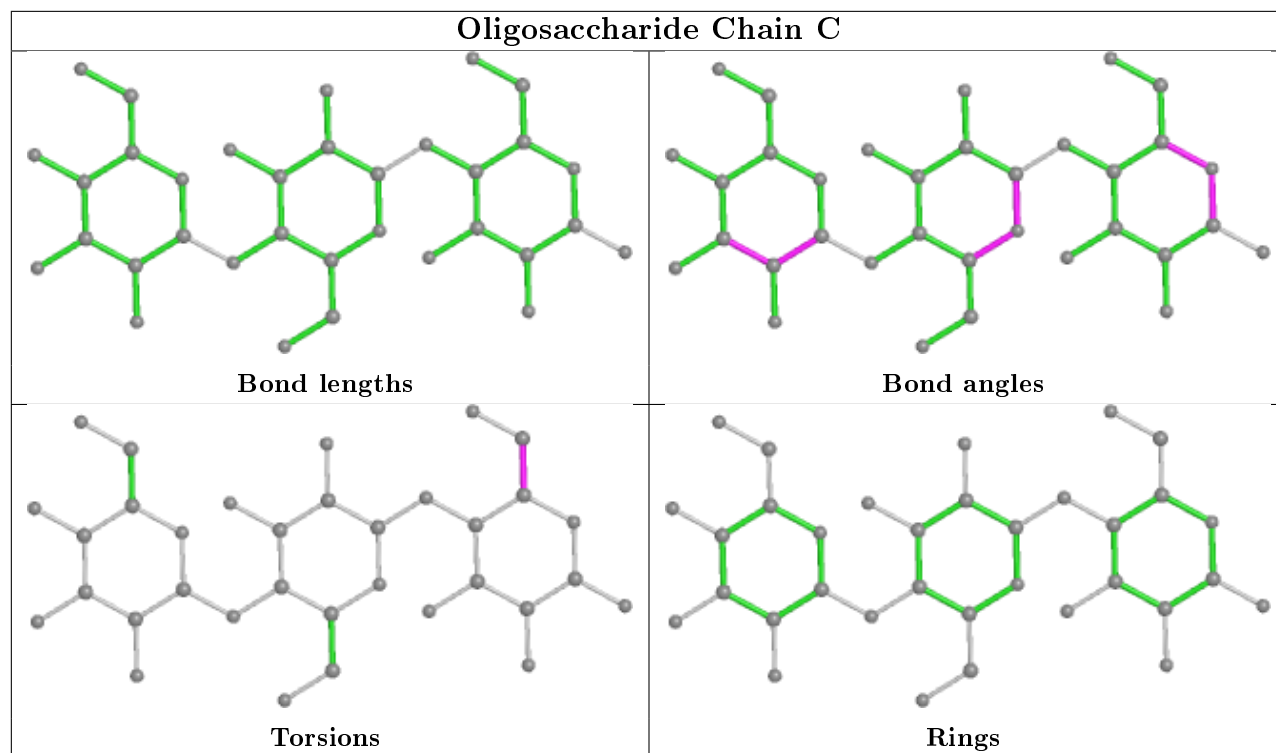
Mol	Chain	Res	Type	Atoms
2	C	2	GLC	O5-C5-C6-O6
3	D	3	GLC	O5-C5-C6-O6
2	B	2	GLC	C4-C5-C6-O6
3	D	3	GLC	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.





5.6 Ligand geometry

Of 6 ligands modelled in this entry, 5 are monoatomic - leaving 1 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	A	1106	-	6,6,6	0.62	0	5,5,5	0.98	1 (20%)

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	A	1106	-	-	3/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	1106	PEG	O2-C2-C1	2.01	118.91	110.07

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	1106	PEG	O1-C1-C2-O2
7	A	1106	PEG	O2-C3-C4-O4
7	A	1106	PEG	C4-C3-O2-C2

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1106	PEG	4	0

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, 95th 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 ²)	Q<0.9
1	A	1030/1084 (95%)	-0.56	7 (0%) 87 86	14, 22, 37, 54	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	101	THR	3.8
1	A	1045	ALA	3.6
1	A	98	ASN	2.6
1	A	99	GLY	2.6
1	A	100	GLY	2.6

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](#)

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, 95th 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.

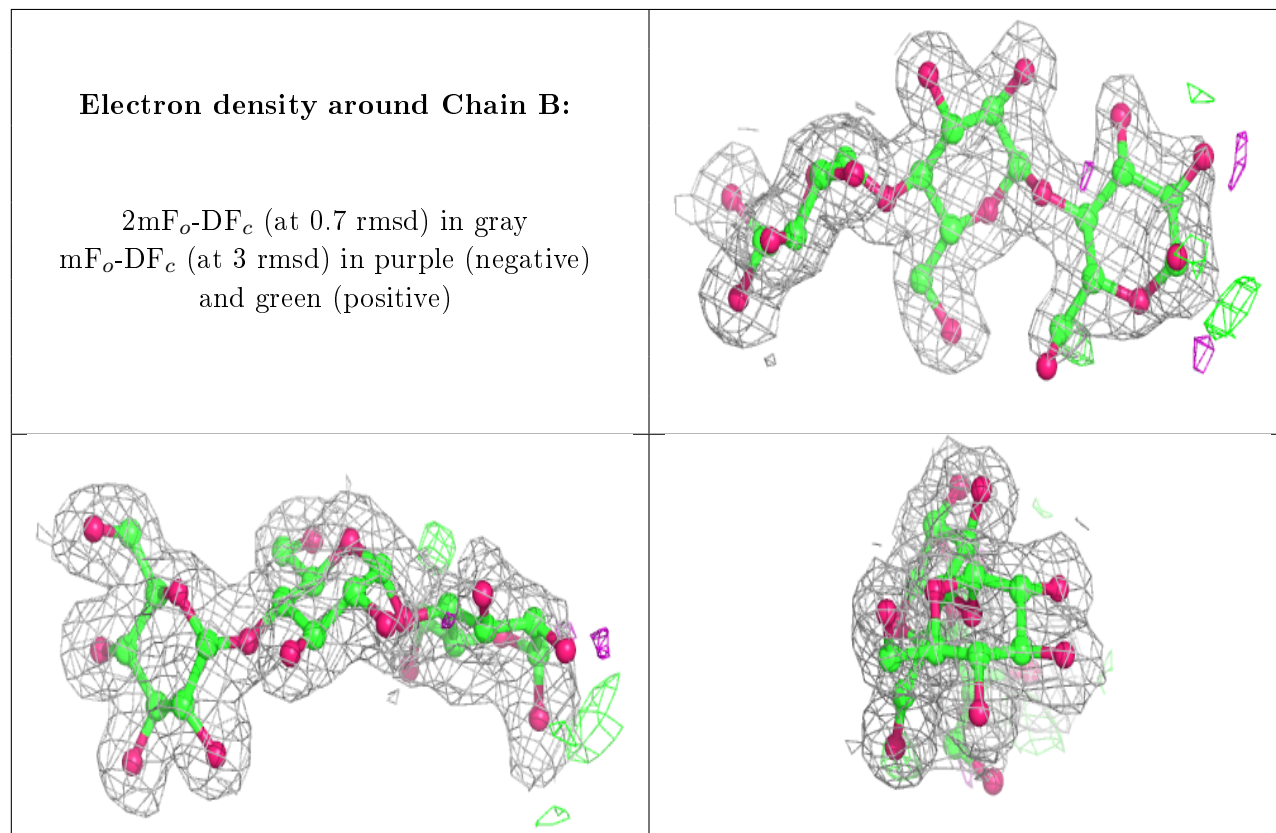
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	GLC	D	2	11/12	0.80	0.27	39,52,56,59	0
2	GLC	C	2	11/12	0.83	0.23	42,52,56,59	0
3	GLC	D	3	11/12	0.85	0.19	61,63,69,69	0
3	GLC	D	1	1/12	0.87	0.19	52,52,52,52	0
2	GLC	C	3	11/12	0.88	0.15	32,38,42,47	0
2	GLC	C	1	1/12	0.91	0.13	54,54,54,54	0
2	GLC	B	2	11/12	0.91	0.20	27,39,47,53	0

Continued on next page...

Continued from previous page...

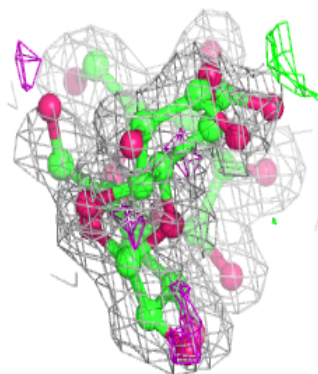
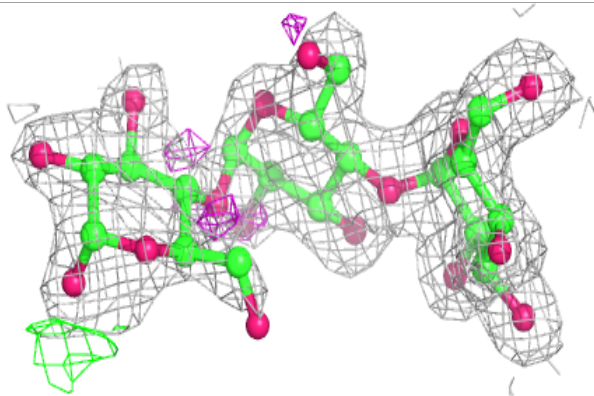
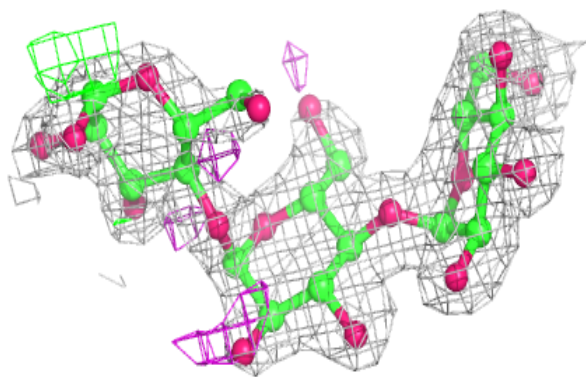
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	GLC	C	4	11/12	0.94	0.08	29,31,35,35	0
2	GLC	B	3	11/12	0.95	0.09	21,23,26,27	0
2	GLC	B	1	1/12	0.96	0.06	38,38,38,38	0
2	GLC	B	4	11/12	0.97	0.05	18,21,23,25	0

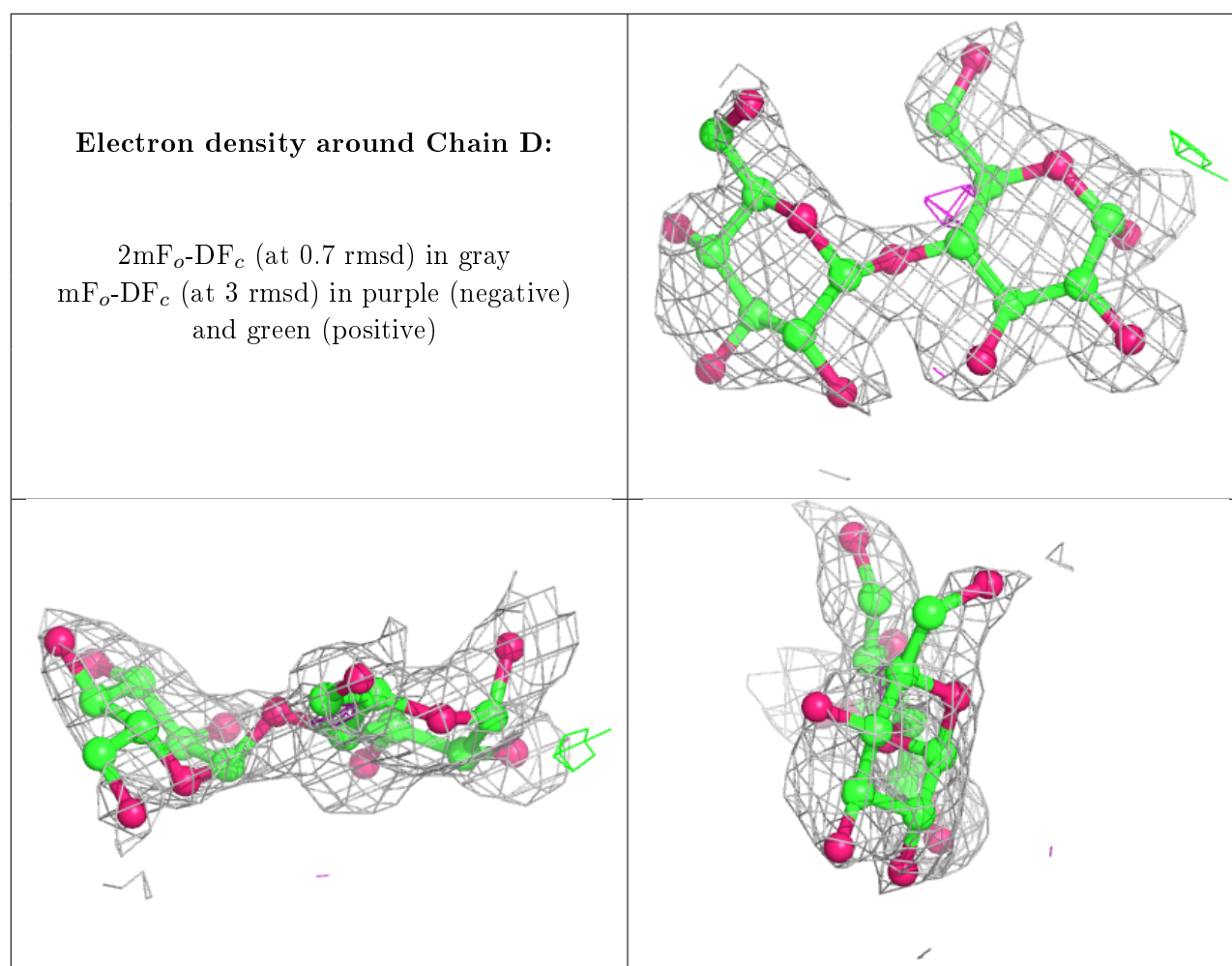
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



Electron density around Chain C:

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





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, 95th 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(\AA^2)	Q<0.9
7	PEG	A	1106	7/7	0.75	0.18	37,41,45,46	0
6	CL	A	1105	1/1	0.90	0.09	48,48,48,48	0
6	CL	A	1104	1/1	0.98	0.05	33,33,33,33	0
4	MG	A	1102	1/1	0.99	0.08	24,24,24,24	0
4	MG	A	1101	1/1	0.99	0.11	24,24,24,24	0
5	CA	A	1103	1/1	1.00	0.04	16,16,16,16	0

6.5 Other polymers [i](#)

There are no such residues in this entry.