wwPDB X-ray Structure Validation Summary Report (i)

Nov 2, 2024 – 12:12 PM EDT

PDB ID	:	6D3U
Title	:	Complex structure of Ulvan lyase from Nonlaben Ulvanivorans- NLR48
Authors	:	Ulaganathan, T.; Cygler, M.
Deposited on	:	2018-04-16
Resolution	:	2.21 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 (i) 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 (1)) were used in the production of this report:

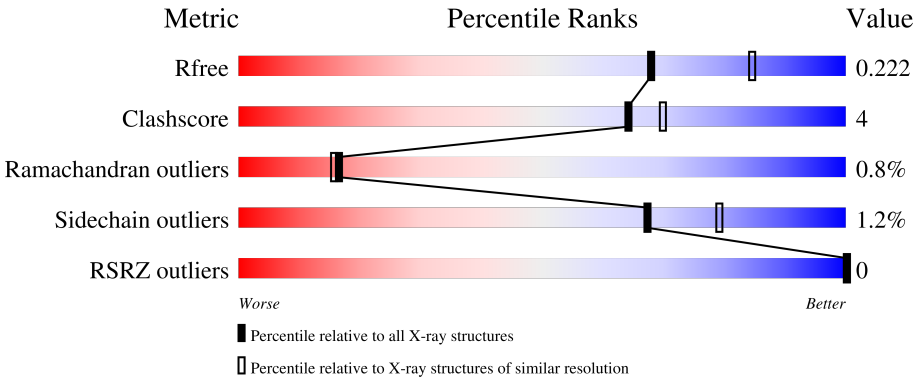
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.21 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	7167 (2.24-2.20)
Clashscore	180529	8096 (2.24-2.20)
Ramachandran outliers	177936	8010 (2.24-2.20)
Sidechain outliers	177891	8011 (2.24-2.20)
RSRZ outliers	164620	7166 (2.24-2.20)

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 >=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 <=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	303	78% 7% • 14%
1	B	303	76% 10% 14%
2	C	4	50% 50%
2	D	4	50% 50%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 4467 atoms, of which 68 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.

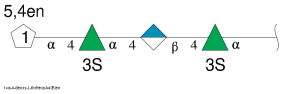
- Molecule 1 is a protein called Ulvan lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	260	Total	C	N	O	S	0	4	0
			2029	1248	366	408	7			
1	B	260	Total	C	N	O	S	0	4	0
			2028	1246	366	409	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	162	MET	LYS	engineered mutation	UNP A0A084JZF2
B	162	MET	LYS	engineered mutation	UNP A0A084JZF2

- Molecule 2 is an oligosaccharide called 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-4)-3-O-sulfo-alpha-L-rhamnopyranose-(1-4)-beta-D-glucopyranuronic acid-(1-4)-3-O-sulfo-alpha-L-rhamnopyranose.

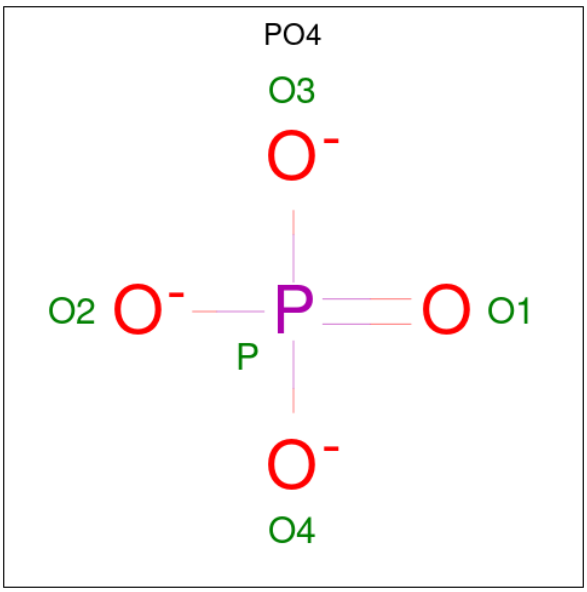


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	4	Total	C	H	O	S	0	0	0
			83	24	31	26	2			
2	D	4	Total	C	H	O	S	0	0	0
			83	24	31	26	2			

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

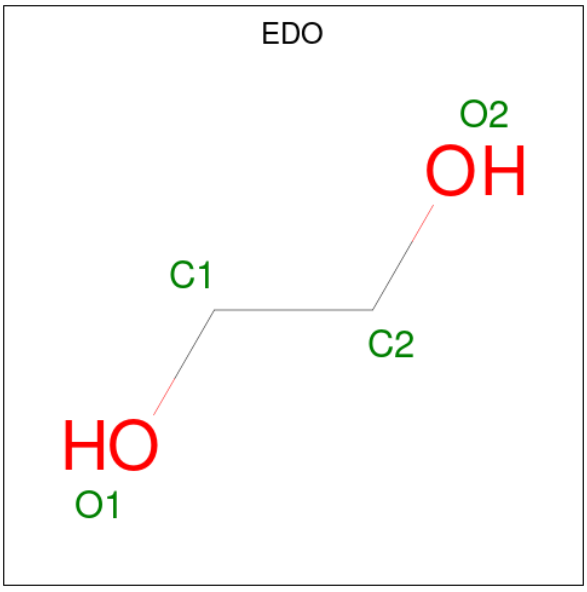
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		
3	B	1	Total	Ca	0	0
			1	1		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



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

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 6 is water.

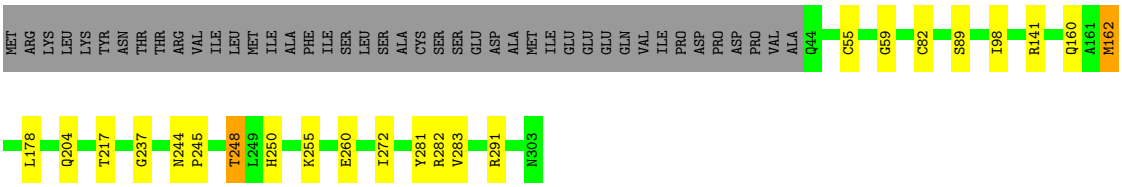
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	103	Total	O	0	0
			103	103		
6	B	109	Total	O	0	0
			109	109		

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.

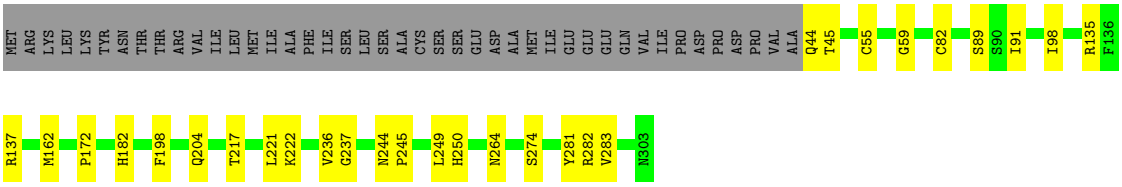
• Molecule 1: Ulvan lyase

Chain A:



• Molecule 1: Ulvan lyase

Chain B:



• Molecule 2: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-4)-3-O-sulfo-alpha-L-rhamnopyranose-(1-4)-beta-D-glucopyranuronic acid-(1-4)-3-O-sulfo-alpha-L-rhamnopyranose

Chain C:

• Molecule 2: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-4)-3-O-sulfo-alpha-L-rhamnopyranose-(1-4)-beta-D-glucopyranuronic acid-(1-4)-3-O-sulfo-alpha-L-rhamnopyranose

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	82.46Å 102.66Å 103.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.05 – 2.21 46.05 – 2.21	Depositor EDS
% Data completeness (in resolution range)	99.8 (46.05-2.21) 99.3 (46.05-2.21)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	1.31 (at 2.20Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.190 , 0.220 0.194 , 0.222	Depositor DCC
R_{free} test set	42394 reflections (4.50%)	wwPDB-VP
Wilson B-factor (Å^2)	49.6	Xtriage
Anisotropy	0.118	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 28.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	0.417 for -h,l,k	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4467	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: PO4, 83Y, BDP, GCD, EDO, CA

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	0.45	0/2081	0.65	0/2819
1	B	0.45	0/2079	0.65	0/2816
All	All	0.45	0/4160	0.65	0/5635

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

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	2029	0	1887	12	0
1	B	2028	0	1878	15	0
2	C	52	31	11	1	0
2	D	52	31	11	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	10	0	0	0	0
4	B	10	0	0	0	0
5	B	4	6	6	1	0
6	A	103	0	0	1	0
6	B	109	0	0	1	0
All	All	4399	68	3793	30	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 30 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:44:GLN:C	1:B:91:ILE:HD11	2.06	0.76
1:A:248:THR:HG21	6:A:589:HOH:O	1.89	0.71
1:A:55[B]:CYS:HB3	1:A:82:CYS:SG	2.32	0.69
1:A:255:LYS:HG2	1:A:260:GLU:HG2	1.78	0.65
1:B:59:GLY:O	1:B:82:CYS:HA	1.98	0.62

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	262/303 (86%)	249 (95%)	11 (4%)	2 (1%)	16	16
1	B	262/303 (86%)	251 (96%)	9 (3%)	2 (1%)	16	16
All	All	524/606 (86%)	500 (95%)	20 (4%)	4 (1%)	16	16

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	245	PRO
1	A	245	PRO
1	B	283	VAL
1	A	283	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	A	214/252 (85%)	211 (99%)	3 (1%)	62 75
1	B	213/252 (84%)	211 (99%)	2 (1%)	75 85
All	All	427/504 (85%)	422 (99%)	5 (1%)	67 79

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

Mol	Chain	Res	Type
1	A	162	MET
1	A	244	ASN
1	A	248	THR
1	B	162	MET
1	B	244	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	264	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

8 monosaccharides 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	$\#\|Z\| > 2$
2	83Y	C	1	2	15,15,15	1.94	6 (40%)	21,23,23	1.33	4 (19%)
2	BDP	C	2	2	12,12,13	1.55	4 (33%)	14,17,19	1.26	1 (7%)
2	83Y	C	3	2	14,14,15	0.93	1 (7%)	18,21,23	1.48	3 (16%)
2	GCD	C	4	2	10,11,12	2.41	2 (20%)	12,15,17	3.08	5 (41%)
2	83Y	D	1	2	15,15,15	1.90	4 (26%)	21,23,23	1.36	3 (14%)
2	BDP	D	2	2	12,12,13	1.57	4 (33%)	14,17,19	1.24	1 (7%)
2	83Y	D	3	2	14,14,15	0.83	0	18,21,23	1.73	6 (33%)
2	GCD	D	4	2	10,11,12	2.40	1 (10%)	12,15,17	2.97	4 (33%)

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	83Y	C	1	2	-	0/5/25/25	0/1/1/1
2	BDP	C	2	2	-	0/4/21/24	0/1/1/1
2	83Y	C	3	2	-	3/5/22/25	0/1/1/1
2	GCD	C	4	2	-	4/4/17/20	0/1/1/1
2	83Y	D	1	2	-	0/5/25/25	0/1/1/1
2	BDP	D	2	2	-	0/4/21/24	0/1/1/1
2	83Y	D	3	2	-	3/5/22/25	0/1/1/1
2	GCD	D	4	2	-	4/4/17/20	0/1/1/1

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	4	GCD	O5-C5	6.94	1.46	1.37
2	D	4	GCD	O5-C5	6.70	1.46	1.37
2	C	1	83Y	O3-S	-4.25	1.44	1.57
2	D	1	83Y	O3-S	-4.16	1.44	1.57
2	D	1	83Y	O3-C3	-3.42	1.38	1.46

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	4	GCD	O5-C5-C4	-7.21	118.32	124.94
2	C	4	GCD	O5-C5-C4	-6.97	118.55	124.94
2	C	4	GCD	O5-C5-C6	6.22	123.29	111.85
2	D	4	GCD	O5-C5-C6	5.07	121.17	111.85
2	C	3	83Y	C2-C3-C4	-3.55	107.82	110.91

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	4	GCD	O5-C5-C6-O6A
2	C	4	GCD	O5-C5-C6-O6B
2	D	3	83Y	C3-O3-S-O1S
2	D	3	83Y	C3-O3-S-O2S
2	D	3	83Y	C3-O3-S-O3S

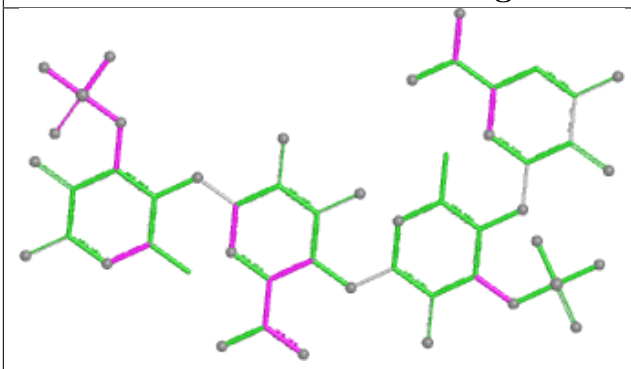
There are no ring outliers.

4 monomers are involved in 3 short contacts:

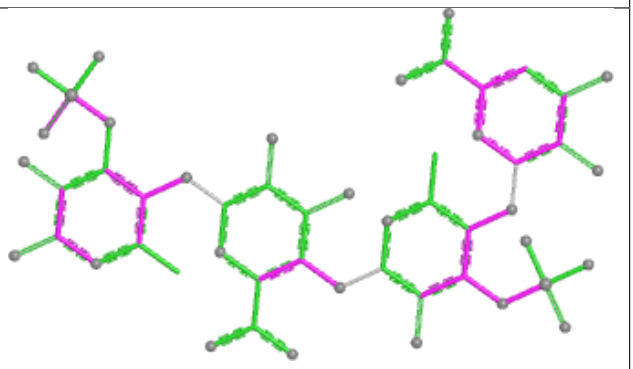
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	3	83Y	1	0
2	D	3	83Y	2	0
2	C	4	GCD	1	0
2	D	4	GCD	2	0

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

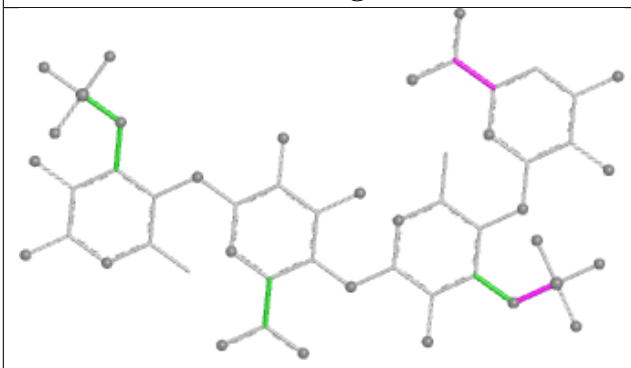
Oligosaccharide Chain C



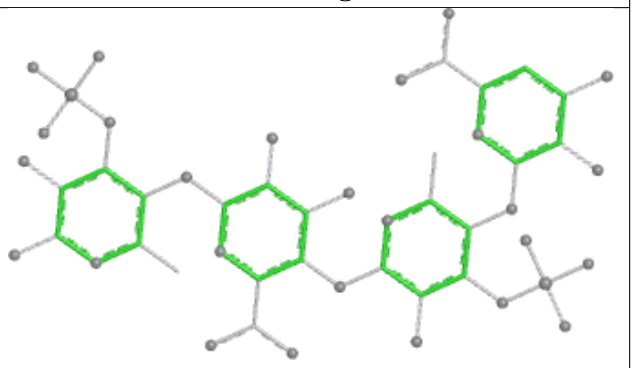
Bond lengths



Bond angles

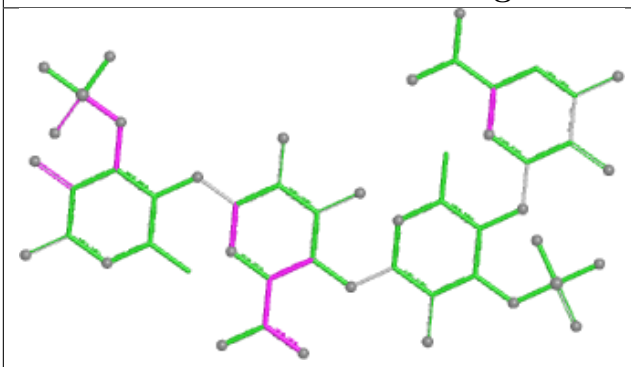


Torsions

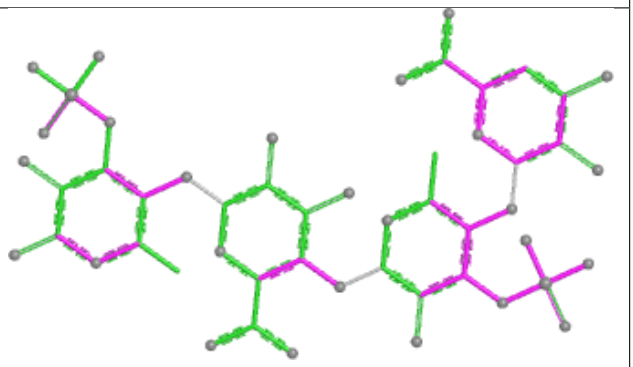


Rings

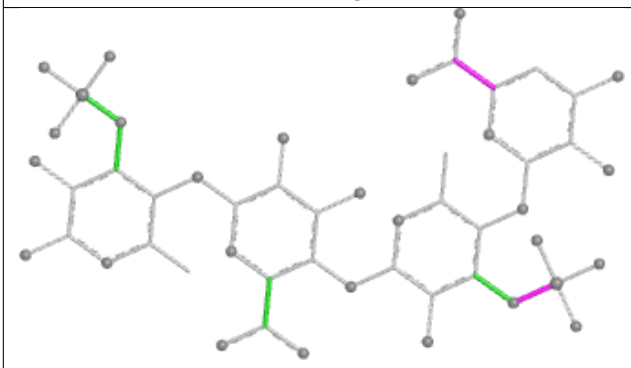
Oligosaccharide Chain D



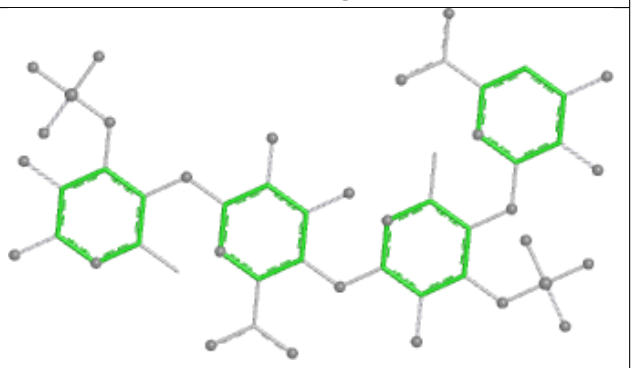
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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	$\#\|Z\| > 2$
4	PO4	A	402	-	4,4,4	1.10	0	6,6,6	0.51	0
5	EDO	B	408	-	3,3,3	0.43	0	2,2,2	0.53	0
4	PO4	B	403	-	4,4,4	0.80	0	6,6,6	0.52	0
4	PO4	A	403	-	4,4,4	0.92	0	6,6,6	0.61	0
4	PO4	B	402	-	4,4,4	0.78	0	6,6,6	0.51	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	EDO	B	408	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	408	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	408	EDO	1	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

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} 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	260/303 (85%)	-1.27	0	100	100	26, 50, 79, 107	4 (1%)
1	B	260/303 (85%)	-1.26	0	100	100	27, 50, 80, 99	4 (1%)
All	All	520/606 (85%)	-1.27	0	100	100	26, 50, 79, 107	8 (1%)

There are no RSRZ outliers to report.

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

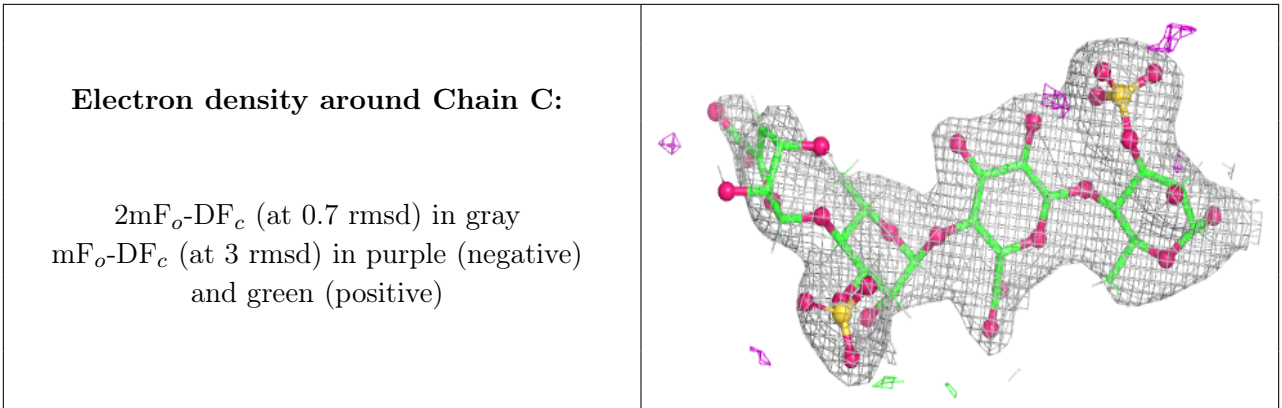
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

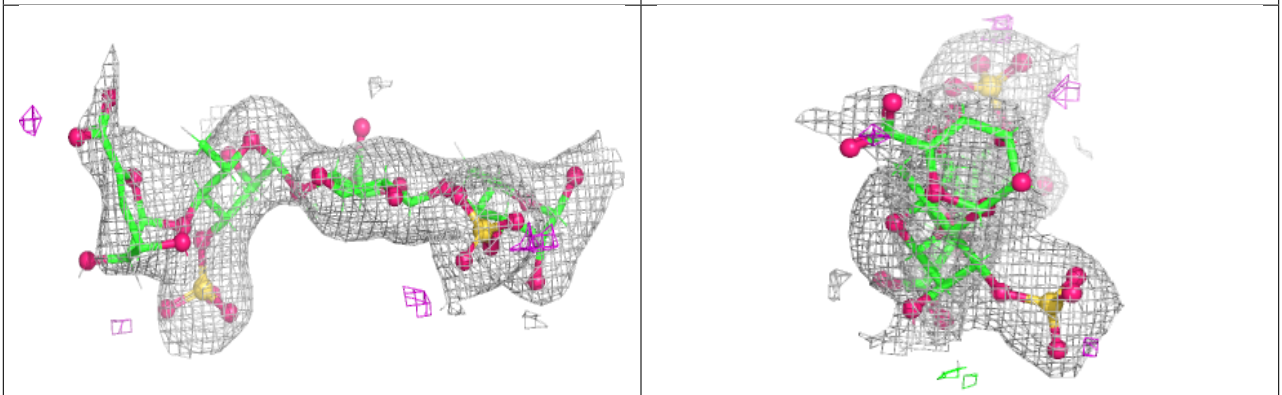
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GCD	D	4	11/12	0.92	0.10	137,149,181,183	0
2	GCD	C	4	11/12	0.94	0.09	127,133,160,160	0
2	83Y	D	3	14/15	0.96	0.08	102,130,155,157	0
2	83Y	C	3	14/15	0.97	0.06	94,121,144,148	0
2	BDP	D	2	12/13	0.97	0.07	69,95,118,125	0
2	BDP	C	2	12/13	0.98	0.05	70,93,111,122	0
2	83Y	D	1	15/15	0.99	0.06	54,65,81,81	0
2	83Y	C	1	15/15	0.99	0.06	54,63,85,85	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:

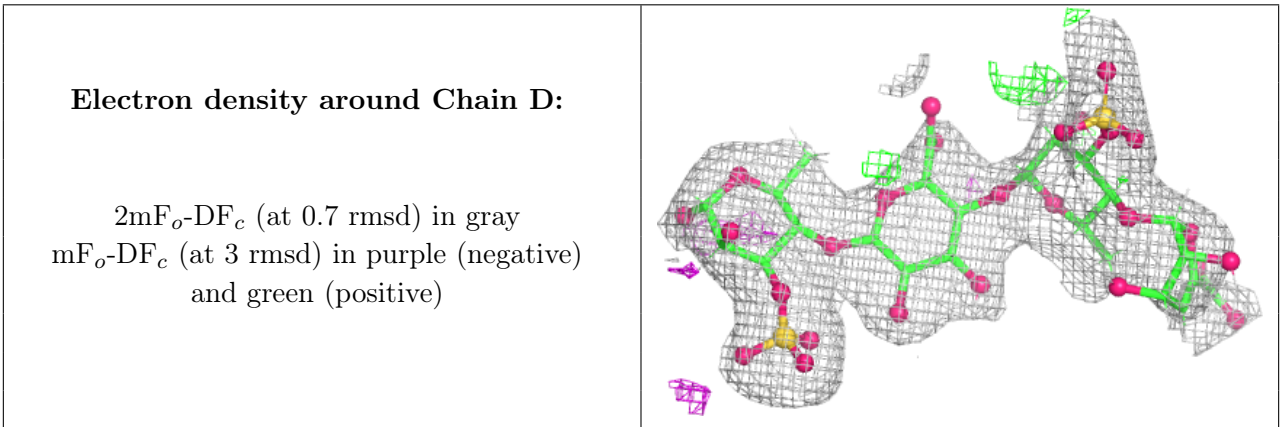
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

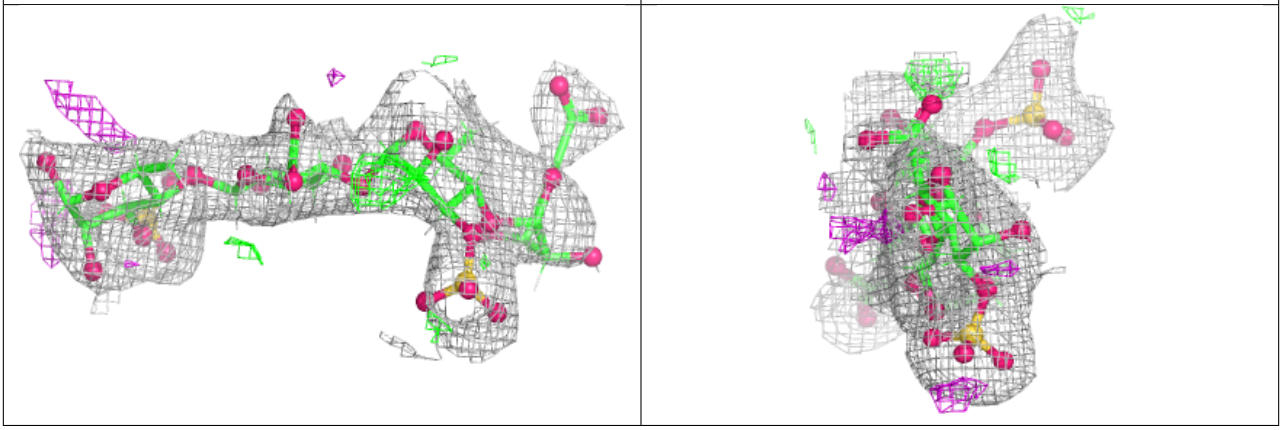




Electron density around Chain D:

$2mF_o$-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, 95^{th} 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
4	PO4	A	403	5/5	0.97	0.06	58,71,85,86	0
4	PO4	B	402	5/5	0.97	0.09	58,65,67,71	5
5	EDO	B	408	4/4	0.97	0.07	61,75,78,93	0
4	PO4	B	403	5/5	0.98	0.05	55,69,79,81	0
4	PO4	A	402	5/5	0.98	0.07	65,65,73,79	5
3	CA	B	401	1/1	0.99	0.02	44,44,44,44	0
3	CA	A	401	1/1	1.00	0.02	42,42,42,42	0

6.5 Other polymers (i)

There are no such residues in this entry.