wwPDB X-ray Structure Validation Summary Report (i)

Mar 4, 2024 – 08:12 PM EST

PDB ID	:	1YNP
Title	:	aldo-keto reductase AKR11C1 from Bacillus halodurans (apo form)
Authors	:	Marquardt, T.; Kostrewa, D.; Winkler, F.K.; Li, X.D.
Deposited on	:	2005-01-25
Resolution	:	1.25 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 (i)) were used in the production of this report:

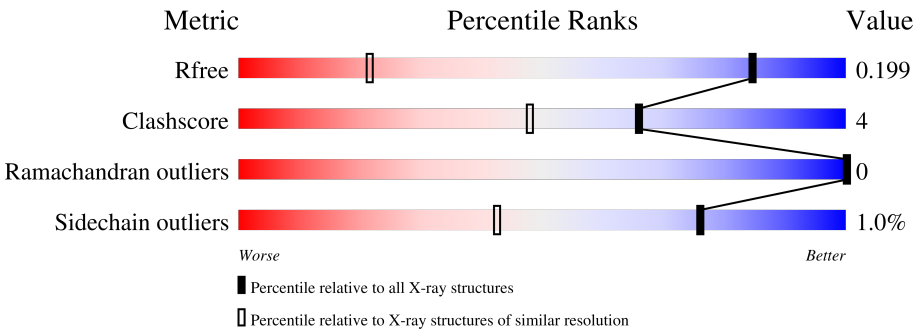
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.25 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	1023 (1.28-1.24)
Clashscore	141614	1060 (1.28-1.24)
Ramachandran outliers	138981	1029 (1.28-1.24)
Sidechain outliers	138945	1028 (1.28-1.24)

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%

Mol	Chain	Length	Quality of chain
1	A	317	80% 9% 11%
1	B	317	86% 7% • 6%
2	C	2	50% 50%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 5368 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 oxidoreductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	283	2279	1432	403	437	7	0	8	0
1	B	298	2413	1520	428	457	8	0	5	0

There are 40 discrepancies between the modelled and reference sequences:

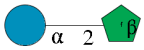
Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	cloning artifact	UNP Q9KE47
A	-19	GLY	-	cloning artifact	UNP Q9KE47
A	-18	SER	-	cloning artifact	UNP Q9KE47
A	-17	SER	-	cloning artifact	UNP Q9KE47
A	-16	HIS	-	cloning artifact	UNP Q9KE47
A	-15	HIS	-	cloning artifact	UNP Q9KE47
A	-14	HIS	-	cloning artifact	UNP Q9KE47
A	-13	HIS	-	cloning artifact	UNP Q9KE47
A	-12	HIS	-	cloning artifact	UNP Q9KE47
A	-11	HIS	-	cloning artifact	UNP Q9KE47
A	-10	SER	-	cloning artifact	UNP Q9KE47
A	-9	SER	-	cloning artifact	UNP Q9KE47
A	-8	GLY	-	cloning artifact	UNP Q9KE47
A	-7	LEU	-	cloning artifact	UNP Q9KE47
A	-6	VAL	-	cloning artifact	UNP Q9KE47
A	-5	PRO	-	cloning artifact	UNP Q9KE47
A	-4	ARG	-	cloning artifact	UNP Q9KE47
A	-3	GLY	-	cloning artifact	UNP Q9KE47
A	-2	SER	-	cloning artifact	UNP Q9KE47
A	-1	HIS	-	cloning artifact	UNP Q9KE47
B	-20	MET	-	cloning artifact	UNP Q9KE47
B	-19	GLY	-	cloning artifact	UNP Q9KE47
B	-18	SER	-	cloning artifact	UNP Q9KE47
B	-17	SER	-	cloning artifact	UNP Q9KE47
B	-16	HIS	-	cloning artifact	UNP Q9KE47

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-15	HIS	-	cloning artifact	UNP Q9KE47
B	-14	HIS	-	cloning artifact	UNP Q9KE47
B	-13	HIS	-	cloning artifact	UNP Q9KE47
B	-12	HIS	-	cloning artifact	UNP Q9KE47
B	-11	HIS	-	cloning artifact	UNP Q9KE47
B	-10	SER	-	cloning artifact	UNP Q9KE47
B	-9	SER	-	cloning artifact	UNP Q9KE47
B	-8	GLY	-	cloning artifact	UNP Q9KE47
B	-7	LEU	-	cloning artifact	UNP Q9KE47
B	-6	VAL	-	cloning artifact	UNP Q9KE47
B	-5	PRO	-	cloning artifact	UNP Q9KE47
B	-4	ARG	-	cloning artifact	UNP Q9KE47
B	-3	GLY	-	cloning artifact	UNP Q9KE47
B	-2	SER	-	cloning artifact	UNP Q9KE47
B	-1	HIS	-	cloning artifact	UNP Q9KE47

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

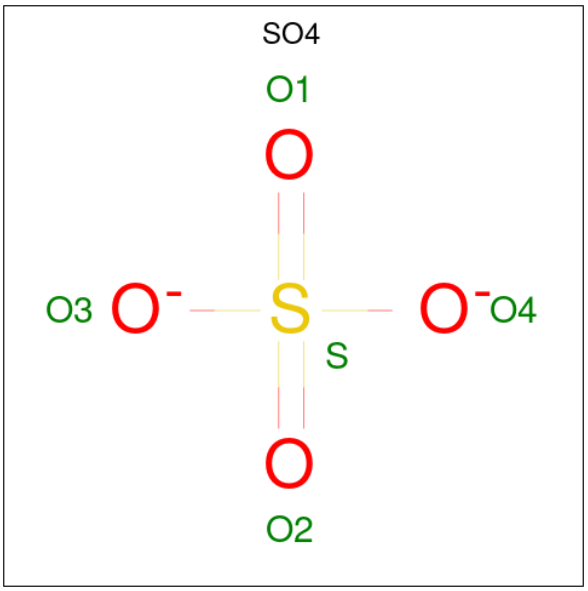


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
2	C	2	23	12	11	0	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

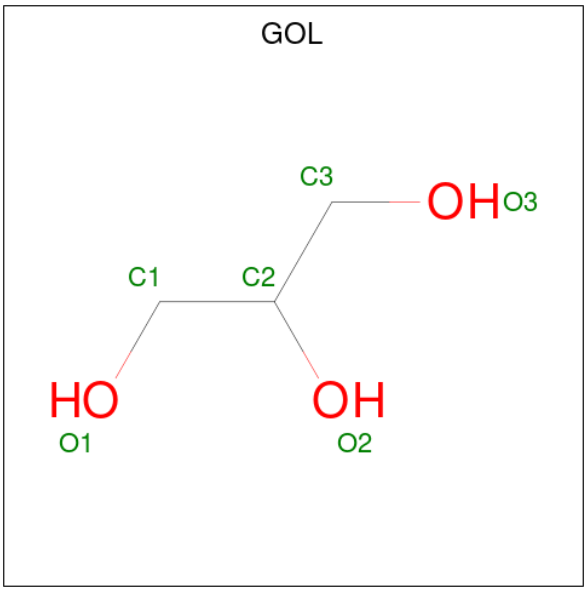
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
3	A	1	1	1	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 6 is water.

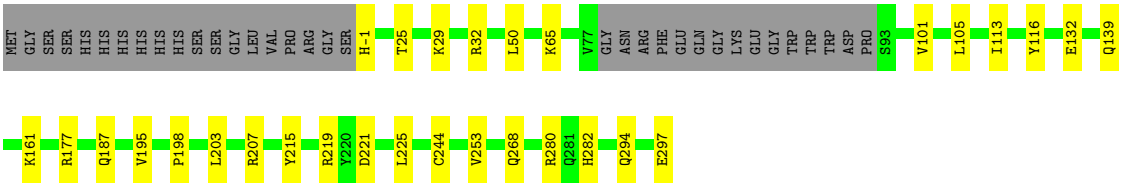
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	283	283	283	0	0
			Total	O		
6	B	303	303	303	0	0

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. 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. 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.

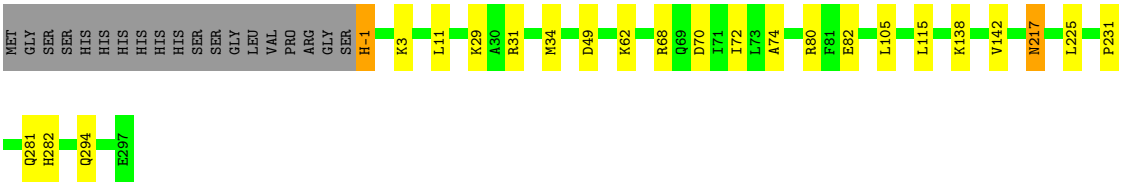
- Molecule 1: oxidoreductase

Chain A:



- Molecule 1: oxidoreductase

Chain B:



- Molecule 2: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	74.67Å 86.57Å 105.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	67.42 – 1.25 66.90 – 1.25	Depositor EDS
% Data completeness (in resolution range)	94.4 (67.42-1.25) 94.4 (66.90-1.25)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.33 (at 1.25Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.170 , 0.206 0.166 , 0.199	Depositor DCC
R_{free} test set	8920 reflections (4.73%)	wwPDB-VP
Wilson B-factor (Å^2)	12.9	Xtriage
Anisotropy	0.389	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 39.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5368	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.60% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GLC, NA, FRU, SO4, GOL

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	0.79	2/2361 (0.1%)	0.86	2/3187 (0.1%)
1	B	0.80	1/2484 (0.0%)	0.87	7/3354 (0.2%)
All	All	0.80	3/4845 (0.1%)	0.87	9/6541 (0.1%)

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	B	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	-1	HIS	C-N	13.36	1.64	1.34
1	A	-1	HIS	C-N	8.12	1.52	1.34
1	A	244	CYS	CB-SG	-5.01	1.73	1.81

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	80	ARG	NE-CZ-NH2	-6.61	117.00	120.30
1	B	-1	HIS	O-C-N	-6.42	112.44	122.70
1	A	177	ARG	NE-CZ-NH1	6.32	123.46	120.30
1	A	177	ARG	NE-CZ-NH2	-6.06	117.27	120.30
1	B	3	LYS	CD-CE-NZ	-6.06	97.76	111.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	-1	HIS	Mainchain

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	2279	0	2290	23	0
1	B	2413	0	2411	14	0
2	C	23	0	21	0	0
3	A	1	0	0	0	0
4	A	40	0	0	2	0
4	B	20	0	0	1	0
5	A	6	0	8	1	0
6	A	283	0	0	8	0
6	B	303	0	0	7	0
All	All	5368	0	4730	38	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 38 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:101:VAL:HG23	1:A:113[A]:ILE:HD12	1.38	1.05
1:B:225:LEU:HD12	6:B:597:HOH:O	1.54	1.05
1:A:101:VAL:CG2	1:A:113[A]:ILE:HD12	1.98	0.93
1:B:29:LYS:NZ	6:B:599:HOH:O	2.06	0.88
1:B:29:LYS:HE3	6:B:474:HOH:O	1.74	0.86

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	287/317 (90%)	283 (99%)	4 (1%)	0	100	100
1	B	301/317 (95%)	295 (98%)	6 (2%)	0	100	100
All	All	588/634 (93%)	578 (98%)	10 (2%)	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	252/272 (93%)	250 (99%)	2 (1%)	81	53
1	B	261/272 (96%)	258 (99%)	3 (1%)	73	39
All	All	513/544 (94%)	508 (99%)	5 (1%)	76	42

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

Mol	Chain	Res	Type
1	A	25	THR
1	A	268	GLN
1	B	68	ARG
1	B	217	ASN
1	B	281	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	69	GLN
1	B	187	GLN
1	B	282	HIS
1	B	217	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	281	GLN

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)

2 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	GLC	C	1	2	11,11,12	0.69	0	15,15,17	1.29	1 (6%)
2	FRU	C	2	2	11,12,12	0.82	0	10,18,18	0.62	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
2	GLC	C	1	2	-	0/2/19/22	0/1/1/1
2	FRU	C	2	2	-	0/5/24/24	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	1	GLC	C1-O5-C5	3.72	117.24	112.19

There are no chirality outliers.

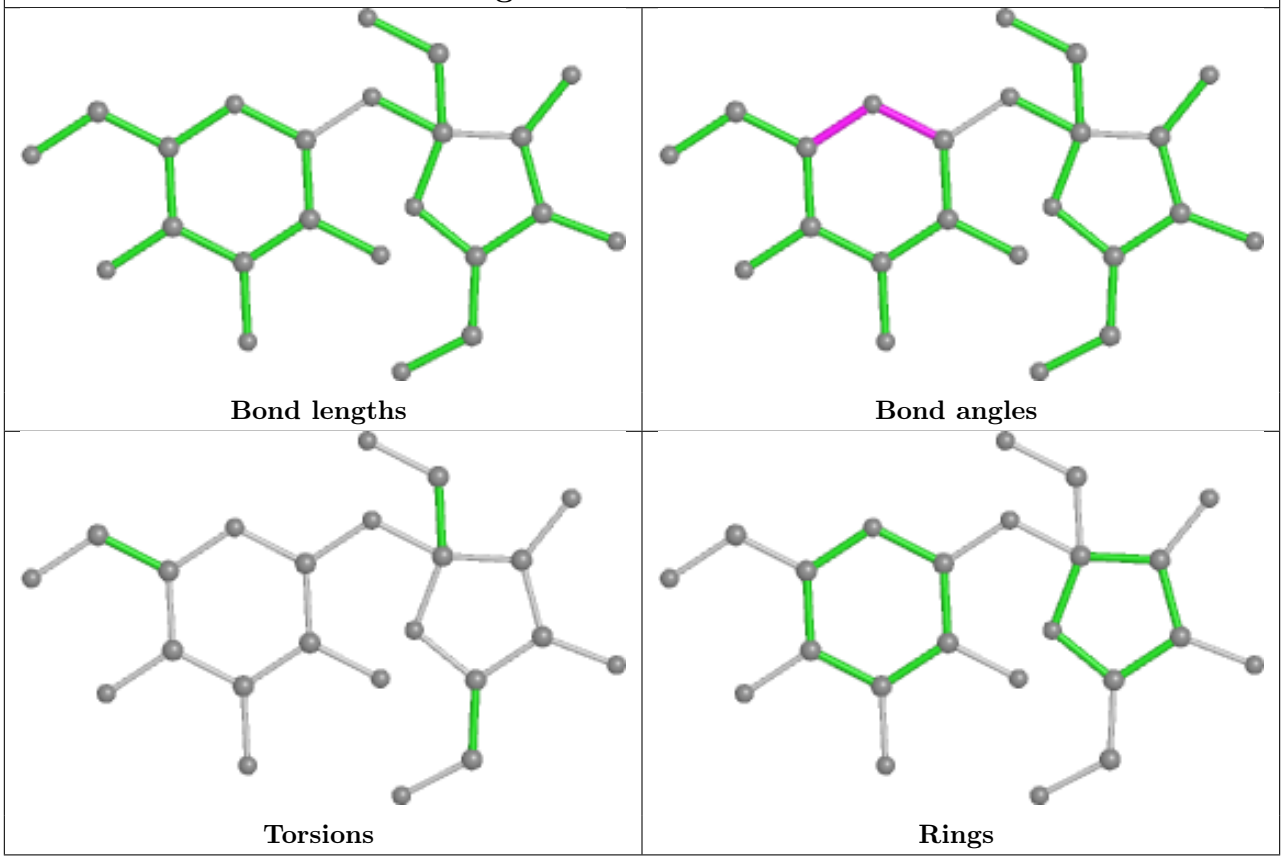
There are no torsion outliers.

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.

Oligosaccharide Chain C



5.6 Ligand geometry (i)

Of 14 ligands modelled in this entry, 1 is monoatomic - leaving 13 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	SO4	A	306	-	4,4,4	0.18	0	6,6,6	0.11	0
4	SO4	B	298	-	4,4,4	0.32	0	6,6,6	0.46	0
4	SO4	A	307	-	4,4,4	0.27	0	6,6,6	0.34	0
5	GOL	A	308	-	5,5,5	0.38	0	5,5,5	1.02	0
4	SO4	A	302	-	4,4,4	0.59	0	6,6,6	0.42	0
4	SO4	B	301	-	4,4,4	0.21	0	6,6,6	0.19	0
4	SO4	A	301	-	4,4,4	0.38	0	6,6,6	1.17	1 (16%)
4	SO4	A	304	-	4,4,4	0.21	0	6,6,6	0.39	0
4	SO4	A	300	-	4,4,4	0.26	0	6,6,6	0.56	0
4	SO4	A	305	-	4,4,4	0.35	0	6,6,6	0.52	0
4	SO4	A	303	-	4,4,4	0.38	0	6,6,6	0.68	0
4	SO4	B	299	-	4,4,4	0.37	0	6,6,6	0.66	0
4	SO4	B	300	-	4,4,4	0.33	0	6,6,6	0.56	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	GOL	A	308	-	-	2/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	301	SO4	O4-S-O3	2.33	119.02	109.06

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	308	GOL	O1-C1-C2-C3
5	A	308	GOL	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	298	SO4	1	0
5	A	308	GOL	1	0
4	A	304	SO4	1	0
4	A	305	SO4	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	-1:HIS	C	1:MET	N	1.64

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

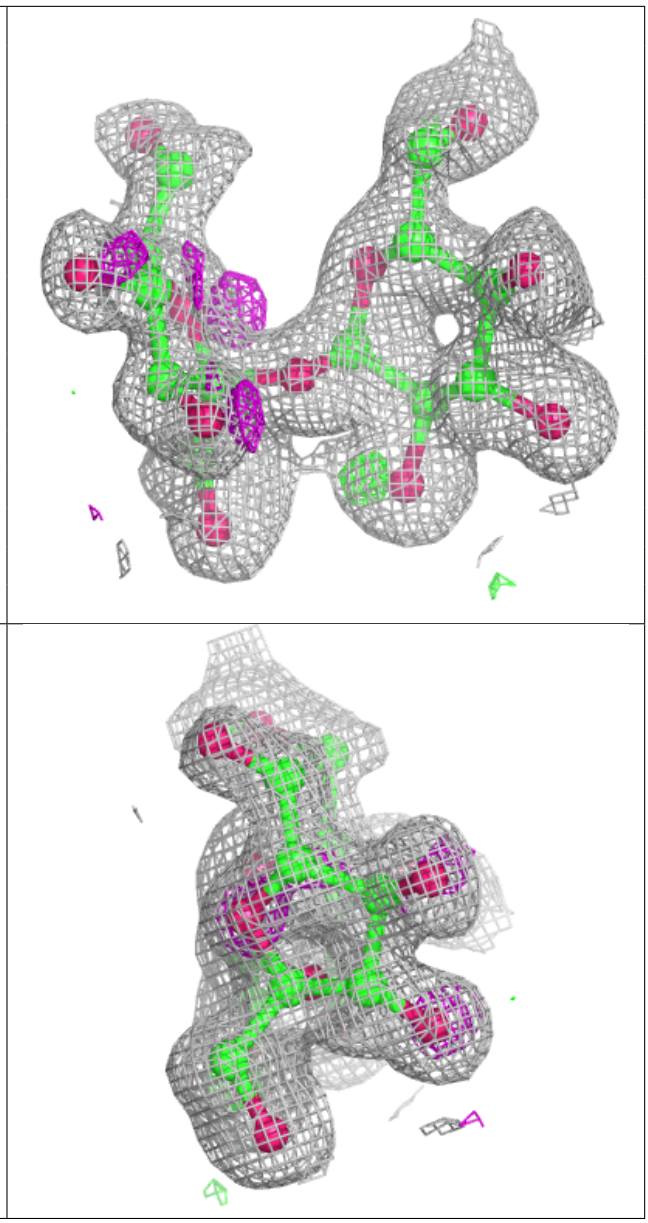
6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

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)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.