wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2023 – 08:19 AM EDT

PDB ID	:	6XSJ
Title	:	X-ray structure of a monoclinic form of alpha amylase from Aspergillus at 1.4 A resolution
Authors	:	McPherson, A.
Deposited on	:	2020-07-15
Resolution	:	1.40 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:

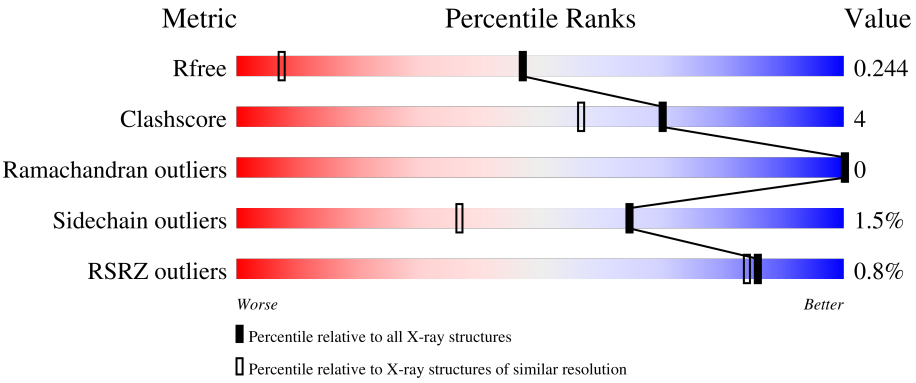
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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

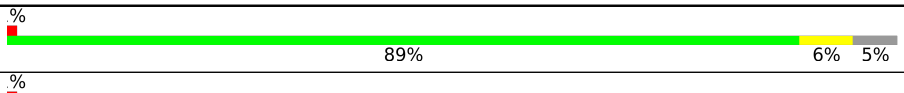
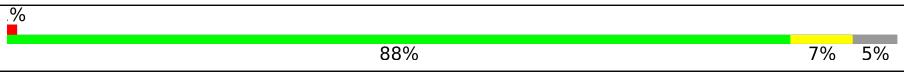
The reported resolution of this entry is 1.40 Å.

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	1714 (1.40-1.40)
Clashscore	141614	1812 (1.40-1.40)
Ramachandran outliers	138981	1763 (1.40-1.40)
Sidechain outliers	138945	1762 (1.40-1.40)
RSRZ outliers	127900	1674 (1.40-1.40)

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	499	
1	B	499	
2	C	2	

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
2	MAN	C	2	-	-	-	X
5	MPD	B	604	-	-	X	-
7	PO4	B	605	-	-	-	X
7	PO4	B	606	-	-	X	X
9	ALA	B	608	-	-	X	X

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 15846 atoms, of which 7086 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 Alpha-amylase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	476	Total	C	H	N	O	S	0	7	0
			7197	2345	3493	597	743	19			
1	B	476	Total	C	H	N	O	S	0	3	0
			7170	2337	3479	596	739	19			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	35	GLN	ARG	conflict	UNP B0FZ76
B	35	GLN	ARG	conflict	UNP B0FZ76

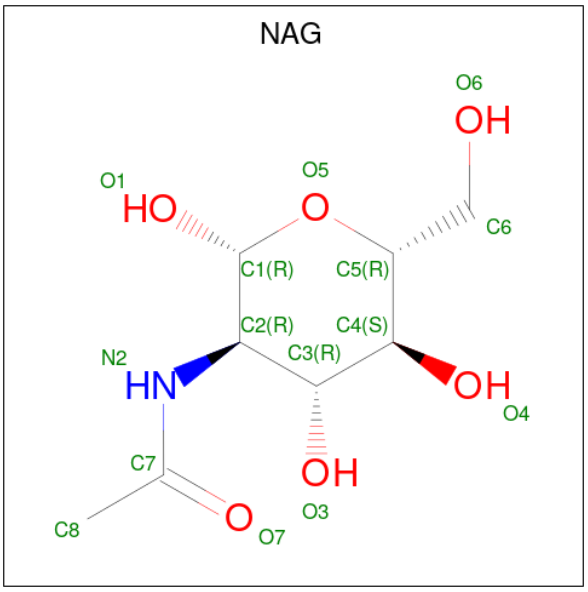
- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	2	Total	C	H	N	O	0	0	0
			49	14	24	1	10			

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

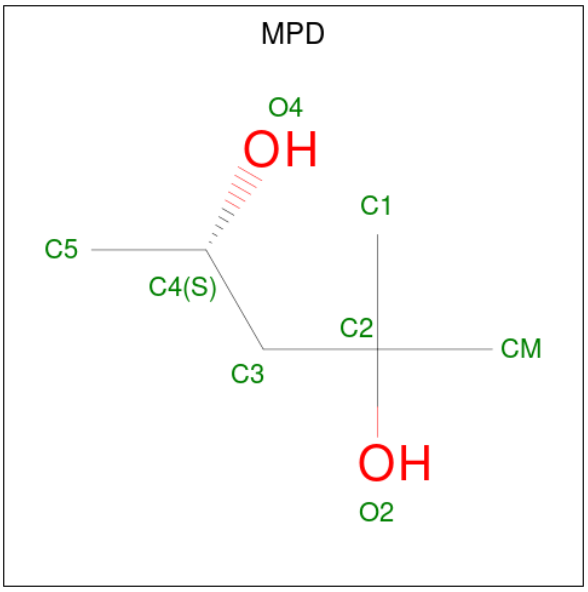
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	1	0
			1	1		
3	B	1	Total	Ca	1	0
			1	1		

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



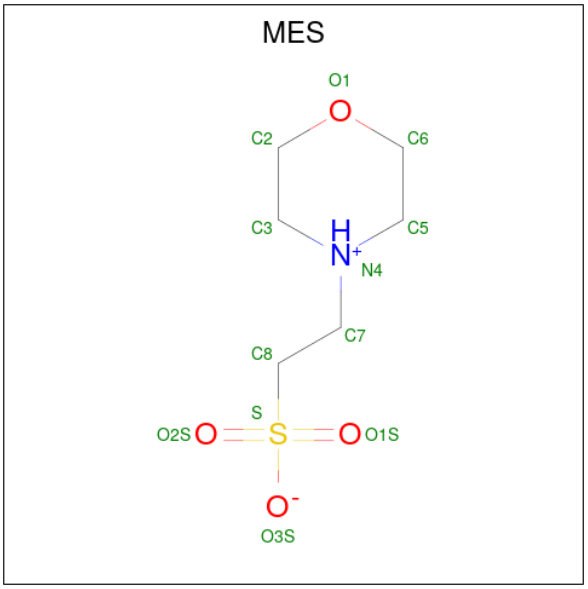
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	0	0
			28	8	14	1	5		

- Molecule 5 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



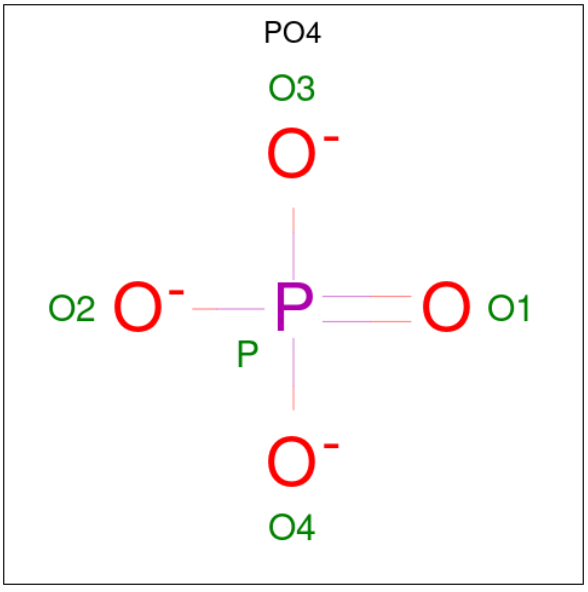
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			22	6	14	2		
5	B	1	Total	C	H	O	0	0
			22	6	14	2		

- Molecule 6 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O	S		
6	A	1	24	6	12	1	4	1	0	0

- Molecule 7 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



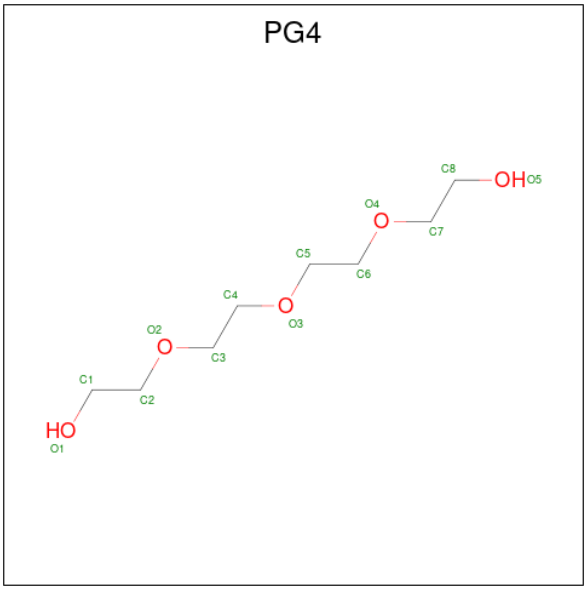
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
7	B	1	5	4	1	0	0

Continued on next page...

Continued from previous page...

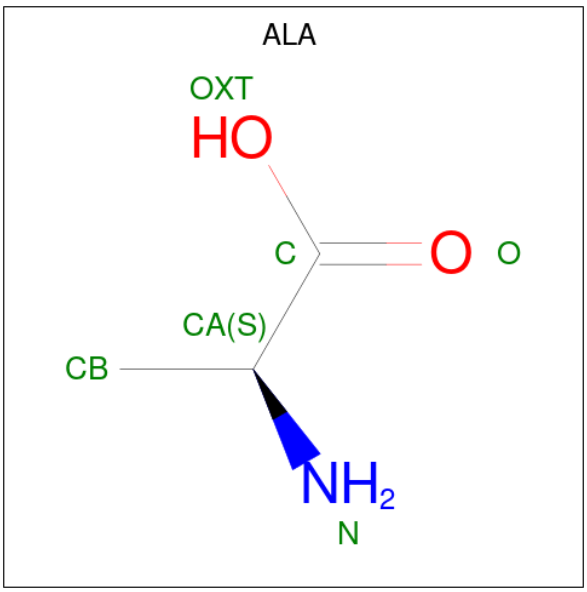
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	1	Total O P 5 4 1	0	0

- Molecule 8 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total C H O 62 16 36 10	0	1

- Molecule 9 is ALANINE (three-letter code: ALA) (formula: $C_3H_7NO_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	B	1	Total C 1 1	0	0

- Molecule 10 is water.

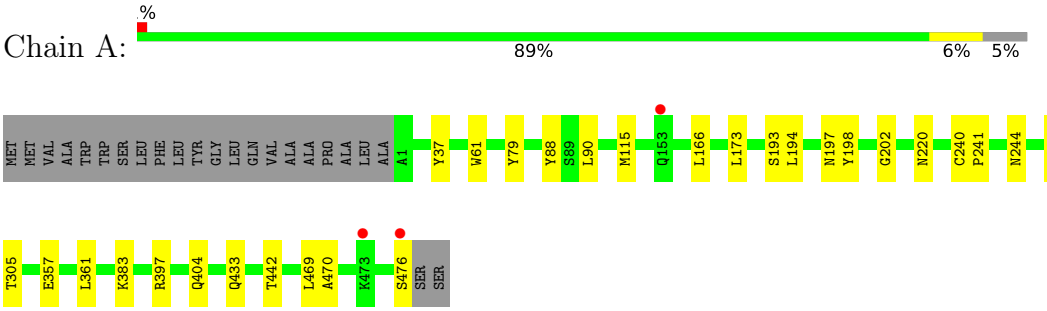
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	794	Total O 795 795	0	1
10	B	464	Total O 464 464	0	0

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.

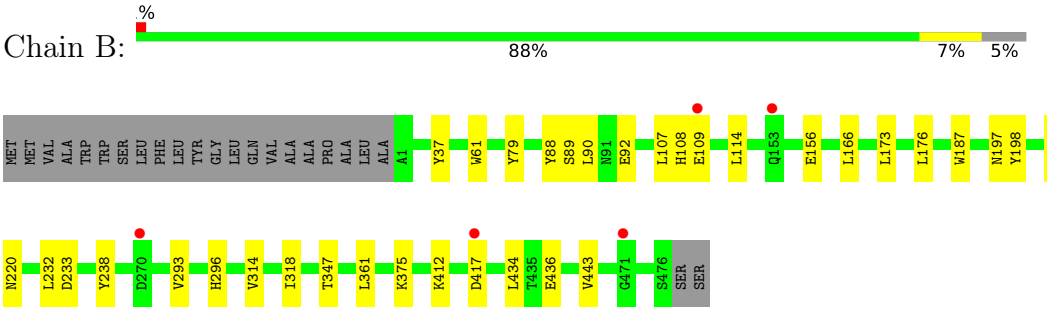
- Molecule 1: Alpha-amylase

Chain A:



- Molecule 1: Alpha-amylase

Chain B:



- Molecule 2: alpha-D-mannopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	64.98Å 95.19Å 74.99Å 90.00° 103.51° 90.00°	Depositor
Resolution (Å)	72.91 – 1.40 72.91 – 1.40	Depositor EDS
% Data completeness (in resolution range)	84.6 (72.91-1.40) 90.1 (72.91-1.40)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	0.15	Depositor
$< I/\sigma(I) >$ [1]	1.08 (at 1.40Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.202 , 0.237 0.210 , 0.244	Depositor DCC
R_{free} test set	7866 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	11.7	Xtriage
Anisotropy	0.257	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 38.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	15846	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.44% 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: MES, MPD, CA, PO4, NAG, PG4, MAN

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	0.64	0/3822	0.68	0/5221
1	B	0.61	0/3797	0.65	0/5186
All	All	0.62	0/7619	0.66	0/10407

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	3704	3493	3485	19	0
1	B	3691	3479	3481	28	0
2	C	25	24	22	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	14	14	13	5	0
5	A	8	14	14	5	0
5	B	8	14	14	6	0
6	A	12	12	12	0	0
7	B	10	0	0	3	0
8	B	26	36	36	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	B	1	0	0	5	0
10	A	795	0	0	6	2
10	B	464	0	0	2	1
All	All	8760	7086	7077	53	2

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 53 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:357:GLU:OE1	10:A:701:HOH:O	2.07	0.71
1:A:470:ALA:O	10:A:702:HOH:O	2.11	0.69
1:A:244:ASN:ND2	10:A:707:HOH:O	2.26	0.68
1:A:90:LEU:HD21	5:A:603:MPD:H52	1.76	0.67
1:B:109:GLU:O	8:B:607[A]:PG4:H52	1.98	0.63

All (2) 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)
10:A:846:HOH:O	10:A:992:HOH:O[2_544]	1.99	0.21
10:A:1467:HOH:O	10:B:1014:HOH:O[1_454]	2.12	0.08

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	481/499 (96%)	470 (98%)	11 (2%)	0	100	100
1	B	477/499 (96%)	466 (98%)	11 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	958/998 (96%)	936 (98%)	22 (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	404/414 (98%)	400 (99%)	4 (1%)	76 53
1	B	400/414 (97%)	392 (98%)	8 (2%)	55 23
All	All	804/828 (97%)	792 (98%)	12 (2%)	65 37

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	220	ASN
1	B	232	LEU
1	B	436	GLU
1	B	296	HIS
1	A	469	LEU

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

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

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	NAG	C	1	1,2	14,14,15	0.96	1 (7%)	17,19,21	0.88	1 (5%)
2	MAN	C	2	2	11,11,12	0.74	0	15,15,17	2.25	4 (26%)

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	NAG	C	1	1,2	-	2/6/23/26	0/1/1/1
2	MAN	C	2	2	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	NAG	C1-C2	2.92	1.56	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	2	MAN	C1-O5-C5	6.32	120.76	112.19
2	C	2	MAN	O2-C2-C3	-3.36	103.41	110.14
2	C	2	MAN	O5-C1-C2	3.22	115.74	110.77
2	C	2	MAN	C1-C2-C3	2.97	113.32	109.67
2	C	1	NAG	O3-C3-C2	2.17	113.96	109.47

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1	NAG	O5-C5-C6-O6
2	C	1	NAG	C4-C5-C6-O6

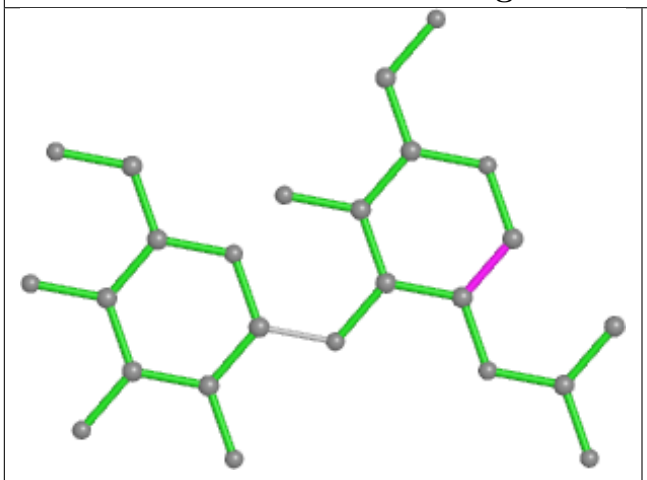
There are no ring outliers.

1 monomer is involved in 2 short contacts:

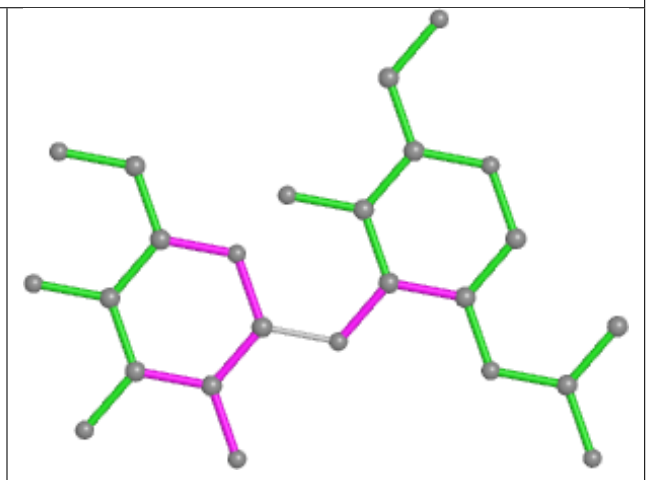
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	NAG	2	0

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

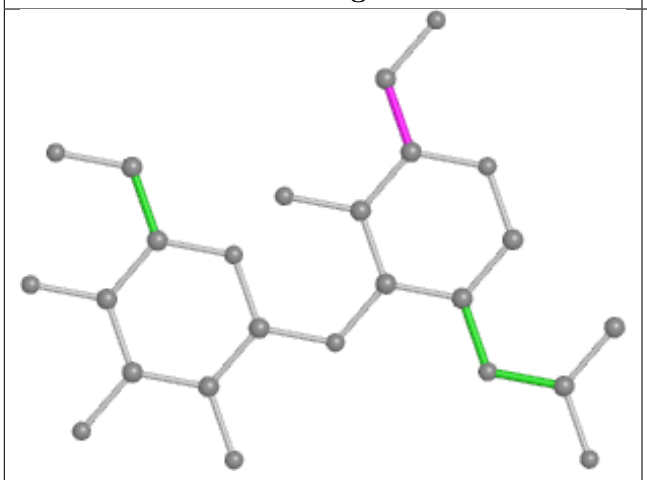
Oligosaccharide Chain C



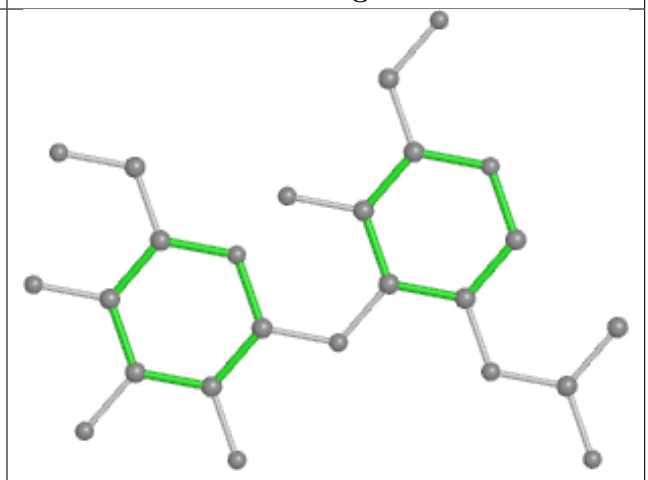
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

Of 11 ligands modelled in this entry, 2 are monoatomic and 1 is modelled with single atom - leaving 8 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	PO4	B	605	-	4,4,4	0.87	0	6,6,6	0.68	0
7	PO4	B	606	-	4,4,4	1.02	0	6,6,6	0.47	0
5	MPD	B	604	-	7,7,7	0.29	0	9,10,10	0.55	0
6	MES	A	604	-	12,12,12	1.90	1 (8%)	14,16,16	2.28	6 (42%)
5	MPD	A	603	-	7,7,7	0.31	0	9,10,10	0.42	0
8	PG4	B	607[B]	-	12,12,12	0.53	0	11,11,11	0.40	0
4	NAG	A	602	1	14,14,15	0.77	1 (7%)	17,19,21	0.64	0
8	PG4	B	607[A]	-	12,12,12	0.53	0	11,11,11	0.44	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	MPD	B	604	-	-	0/5/5/5	-
6	MES	A	604	-	-	1/6/14/14	0/1/1/1
5	MPD	A	603	-	-	2/5/5/5	-
8	PG4	B	607[B]	-	-	7/10/10/10	-
4	NAG	A	602	1	-	2/6/23/26	0/1/1/1
8	PG4	B	607[A]	-	-	7/10/10/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	604	MES	C8-S	-6.38	1.68	1.77
4	A	602	NAG	O5-C1	2.12	1.47	1.43

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	604	MES	C5-N4-C3	6.07	122.49	108.83
6	A	604	MES	C6-C5-N4	2.73	114.24	110.10
6	A	604	MES	O3S-S-C8	2.40	109.65	105.77
6	A	604	MES	C7-N4-C5	2.17	116.78	111.23
6	A	604	MES	O2S-S-C8	2.13	109.48	106.92

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	604	MES	C8-C7-N4-C5
4	A	602	NAG	C4-C5-C6-O6
8	B	607[A]	PG4	C4-C3-O2-C2
8	B	607[B]	PG4	C4-C3-O2-C2
4	A	602	NAG	O5-C5-C6-O6

There are no ring outliers.

7 monomers are involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	605	PO4	1	0
7	B	606	PO4	2	0
5	B	604	MPD	6	0
5	A	603	MPD	5	0
8	B	607[B]	PG4	2	0
4	A	602	NAG	5	0
8	B	607[A]	PG4	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

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	476/499 (95%)	-0.09	3 (0%)	89	88	11, 15, 26, 51	0
1	B	476/499 (95%)	-0.04	5 (1%)	80	79	10, 16, 28, 53	0
All	All	952/998 (95%)	-0.07	8 (0%)	86	84	10, 16, 27, 53	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	471	GLY	3.5
1	A	476	SER	3.1
1	B	153	GLN	2.3
1	B	417	ASP	2.3
1	A	473	LYS	2.2

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

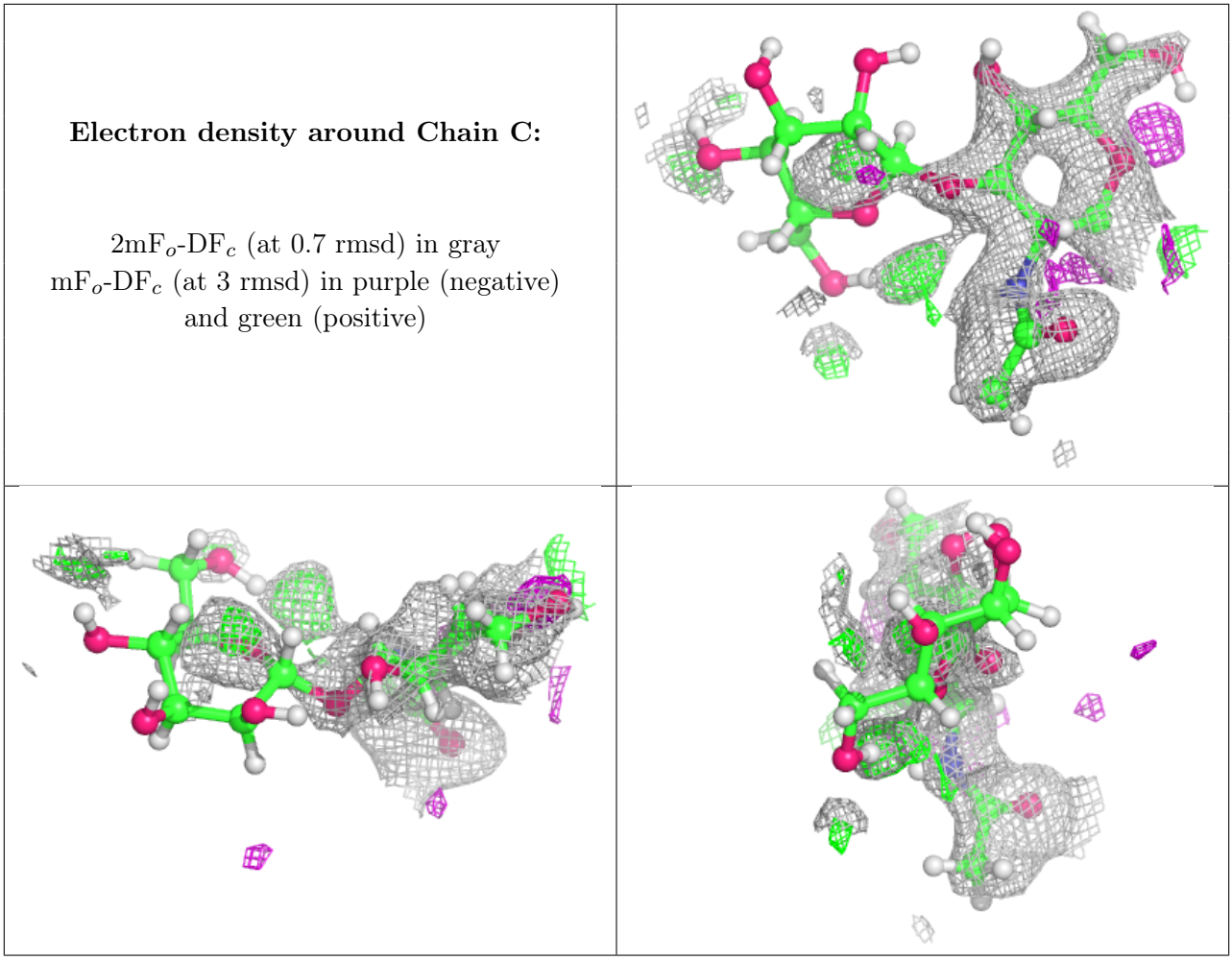
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	MAN	C	2	11/12	0.41	0.77	59,70,83,85	22
2	NAG	C	1	14/15	0.71	0.20	30,43,54,62	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

LIGAND-RSR INFOmissingINFO

6.5 Other polymers

There are no such residues in this entry.