wwPDB X-ray Structure Validation Summary Report (i)

Oct 8, 2020 – 03:06 PM BST

PDB ID	:	6Y6Q
Title	:	Structure of Andes virus envelope glycoprotein Gc in postfusion conformation
Authors	:	Serris, A.; Rey, F.A.; Guardado-Calvo, P.
Deposited on	:	2020-02-27
Resolution	:	2.70 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 following versions of software and data (see references (1)) were used in the production of this report:

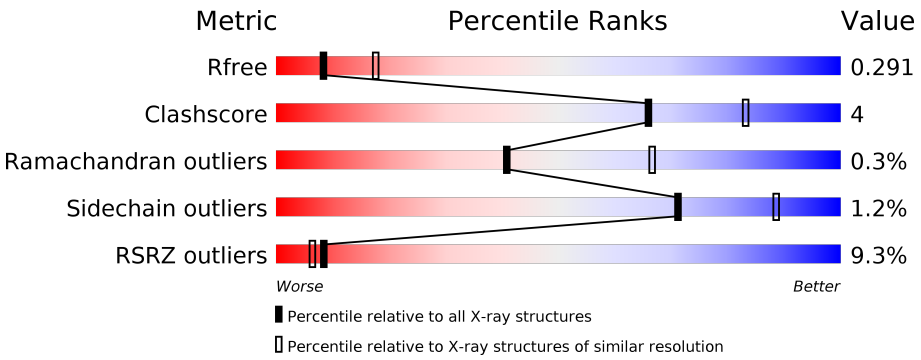
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.6

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 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 <=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	497	7%, 70%, 10%, 20%
2	B	4	50%, 50%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3098 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 Envelope polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	396	3034	1902	512	587	33	0	0	0

There are 41 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	650	ARG	-	expression tag	UNP Q9E006
A	651	SER	-	expression tag	UNP Q9E006
A	1108	GLY	-	expression tag	UNP Q9E006
A	1109	PRO	-	expression tag	UNP Q9E006
A	1110	PHE	-	expression tag	UNP Q9E006
A	1111	GLU	-	expression tag	UNP Q9E006
A	1112	ASP	-	expression tag	UNP Q9E006
A	1113	ASP	-	expression tag	UNP Q9E006
A	1114	ASP	-	expression tag	UNP Q9E006
A	1115	ASP	-	expression tag	UNP Q9E006
A	1116	LYS	-	expression tag	UNP Q9E006
A	1117	ALA	-	expression tag	UNP Q9E006
A	1118	GLY	-	expression tag	UNP Q9E006
A	1119	TRP	-	expression tag	UNP Q9E006
A	1120	SER	-	expression tag	UNP Q9E006
A	1121	HIS	-	expression tag	UNP Q9E006
A	1122	PRO	-	expression tag	UNP Q9E006
A	1123	GLN	-	expression tag	UNP Q9E006
A	1124	PHE	-	expression tag	UNP Q9E006
A	1125	GLU	-	expression tag	UNP Q9E006
A	1126	LYS	-	expression tag	UNP Q9E006
A	1127	GLY	-	expression tag	UNP Q9E006
A	1128	GLY	-	expression tag	UNP Q9E006
A	1129	GLY	-	expression tag	UNP Q9E006
A	1130	SER	-	expression tag	UNP Q9E006
A	1131	GLY	-	expression tag	UNP Q9E006
A	1132	GLY	-	expression tag	UNP Q9E006

Continued on next page...

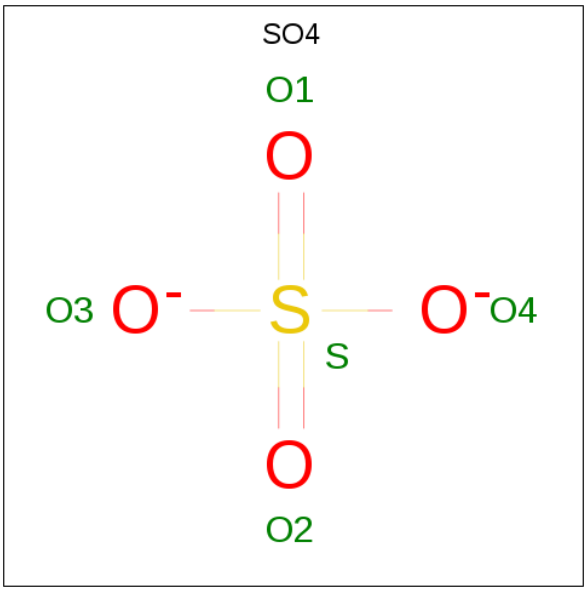
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1133	GLY	-	expression tag	UNP Q9E006
A	1134	SER	-	expression tag	UNP Q9E006
A	1135	GLY	-	expression tag	UNP Q9E006
A	1136	GLY	-	expression tag	UNP Q9E006
A	1137	GLY	-	expression tag	UNP Q9E006
A	1138	SER	-	expression tag	UNP Q9E006
A	1139	TRP	-	expression tag	UNP Q9E006
A	1140	SER	-	expression tag	UNP Q9E006
A	1141	HIS	-	expression tag	UNP Q9E006
A	1142	PRO	-	expression tag	UNP Q9E006
A	1143	GLN	-	expression tag	UNP Q9E006
A	1144	PHE	-	expression tag	UNP Q9E006
A	1145	GLU	-	expression tag	UNP Q9E006
A	1146	LYS	-	expression tag	UNP Q9E006

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	4	50	28	2	20	0	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 4 is water.

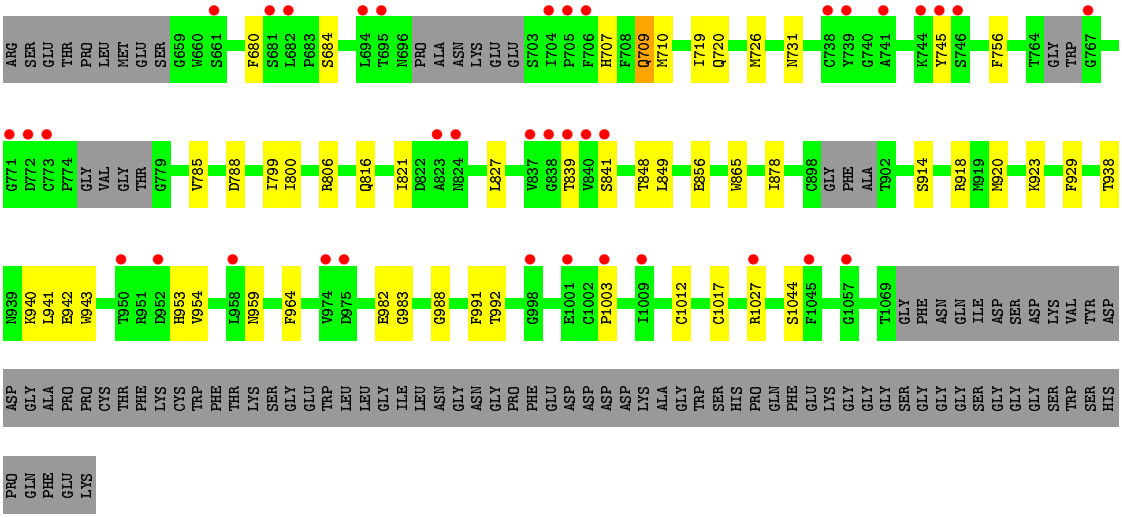
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	4	Total O 4 4	0	0

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: Envelope polyprotein

Chain A:



• Molecule 2: alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	92.65Å 92.65Å 318.55Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.80 – 2.70 39.80 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.6 (39.80-2.70) 99.5 (39.80-2.70)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.75 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.14rc3	Depositor
R, R_{free}	0.234 , 0.287 0.242 , 0.291	Depositor DCC
R_{free} test set	719 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å^2)	67.5	Xtriage
Anisotropy	0.795	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 38.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	3098	wwPDB-VP
Average B, all atoms (Å^2)	98.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% 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: SO4, BMA, NAG, 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	0.26	0/3102	0.47	0/4199

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 (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	3034	0	2917	24	1
2	B	50	0	43	0	0
3	A	10	0	0	0	0
4	A	4	0	0	0	0
All	All	3098	0	2960	24	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 4.

The worst 5 of 24 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:731:ASN:ND2	1:A:788:ASP:OD1	2.30	0.63
1:A:940:LYS:HE2	1:A:942:GLU:HG3	1.82	0.60
1:A:710:MET:HE1	1:A:954:VAL:HG21	1.86	0.56
1:A:684:SER:HA	1:A:806:ARG:NH2	2.24	0.53
1:A:914:SER:O	1:A:918:ARG:HG3	2.09	0.52

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:A:745:TYR:O	1:A:816:GLN:NE2[17_554]	2.10	0.10

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	386/497 (78%)	364 (94%)	21 (5%)	1 (0%)	41 66

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	841	SER

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	345/422 (82%)	341 (99%)	4 (1%)	71 88

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	709	GLN
1	A	839	THR
1	A	938	THR
1	A	1044	SER

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

Mol	Chain	Res	Type
1	A	709	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)

4 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	B	1	1,2	14,14,15	0.31	0	17,19,21	0.62	1 (5%)
2	NAG	B	2	2	14,14,15	0.40	0	17,19,21	0.47	0
2	BMA	B	3	2	11,11,12	0.65	0	15,15,17	0.84	0
2	MAN	B	4	2	11,11,12	0.72	0	15,15,17	1.07	2 (13%)

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	B	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	B	2	2	-	0/6/23/26	0/1/1/1
2	BMA	B	3	2	-	0/2/19/22	0/1/1/1
2	MAN	B	4	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	4	MAN	C1-O5-C5	2.72	115.88	112.19
2	B	4	MAN	O2-C2-C3	-2.23	105.66	110.14
2	B	1	NAG	C1-O5-C5	2.12	115.06	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

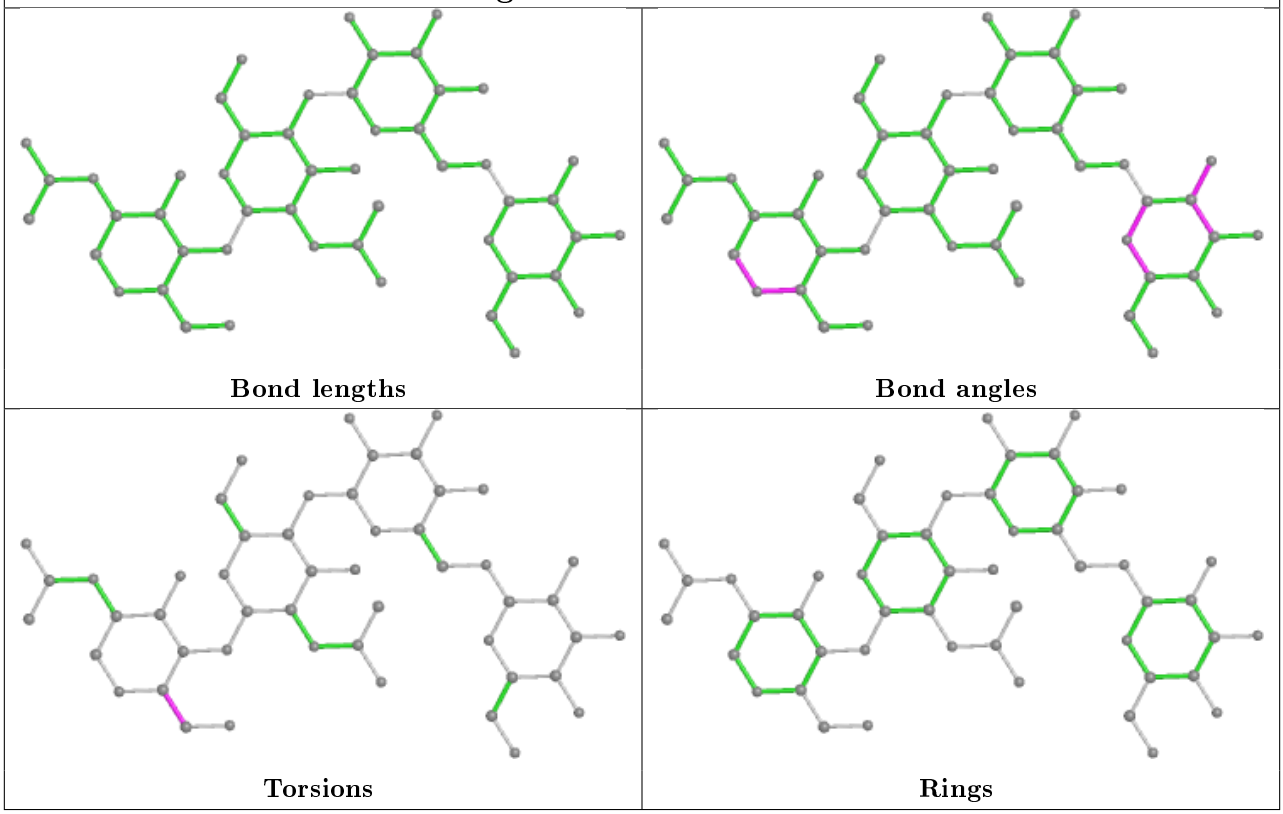
Mol	Chain	Res	Type	Atoms
2	B	1	NAG	O5-C5-C6-O6
2	B	1	NAG	C4-C5-C6-O6

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.

Oligosaccharide Chain B



5.6 Ligand geometry (i)

2 ligands 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	$\#\|Z\| > 2$
3	SO4	A	1206	-	4,4,4	0.15	0	6,6,6	0.07	0
3	SO4	A	1205	-	4,4,4	0.14	0	6,6,6	7.23	3 (50%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1205	SO4	O4-S-O1	-15.62	27.77	109.31

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1205	SO4	O4-S-O3	7.84	142.51	109.06
3	A	1205	SO4	O2-S-O1	2.19	125.61	109.43

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	396/497 (79%)	0.47	37 (9%) 8 6	61, 93, 149, 254	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	838	GLY	9.8
1	A	772	ASP	9.4
1	A	839	THR	9.2
1	A	841	SER	5.6
1	A	706	PHE	4.9

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)

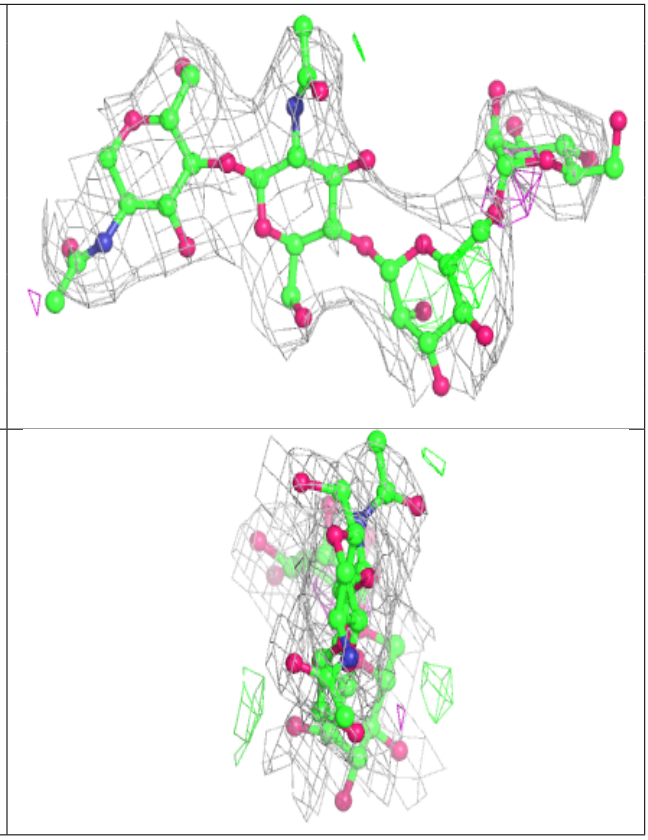
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	BMA	B	3	11/12	0.55	0.19	113,120,126,135	0
2	MAN	B	4	11/12	0.73	0.35	128,132,138,138	0
2	NAG	B	2	14/15	0.90	0.15	83,99,114,114	0
2	NAG	B	1	14/15	0.94	0.12	57,77,86,89	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 B:

$2mF_o$-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(Å^2)	Q<0.9
3	SO4	A	1206	5/5	0.86	0.16	126,128,131,134	0
3	SO4	A	1205	5/5	0.94	0.38	174,174,174,174	5

6.5 Other polymers (i)

There are no such residues in this entry.