wwPDB X-ray Structure Validation Summary Report (i)

Oct 8, 2020 – 03:07 PM BST

PDB ID	:	6Y68
Title	:	Structure of Maporal virus envelope glycoprotein Gc in postfusion conformation
Authors	:	Serris, A.; Rey, F.A.; Guardado-Calvo, P.
Deposited on	:	2020-02-26
Resolution	:	2.40 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 following versions of software and data (see references (i)) were used in the production of this report:

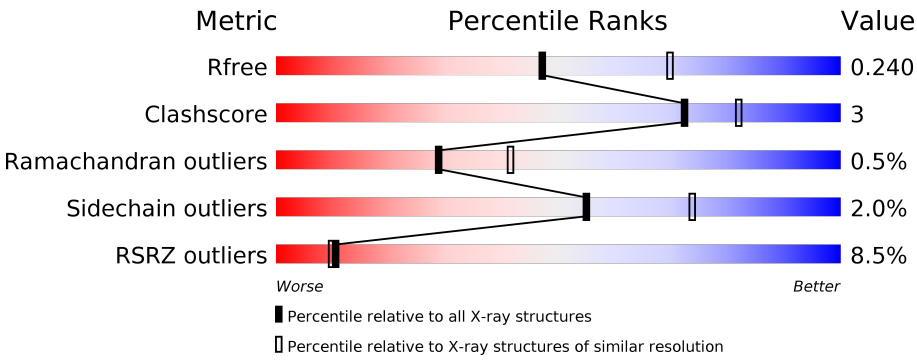
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 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	494	7%, 76%, 7%, •, 16%
2	B	4	25%, 75%

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	B	4	-	-	-	X
3	CD	A	1208	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3291 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	413	3154	1975	533	612	34	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1108	PRO	-	expression tag	UNP Q5MYC0
A	1109	PHE	-	expression tag	UNP Q5MYC0
A	1110	GLU	-	expression tag	UNP Q5MYC0
A	1111	ASP	-	expression tag	UNP Q5MYC0
A	1112	ASP	-	expression tag	UNP Q5MYC0
A	1113	ASP	-	expression tag	UNP Q5MYC0
A	1114	ASP	-	expression tag	UNP Q5MYC0
A	1115	LYS	-	expression tag	UNP Q5MYC0
A	1116	ALA	-	expression tag	UNP Q5MYC0
A	1117	GLY	-	expression tag	UNP Q5MYC0
A	1118	TRP	-	expression tag	UNP Q5MYC0
A	1119	SER	-	expression tag	UNP Q5MYC0
A	1120	HIS	-	expression tag	UNP Q5MYC0
A	1121	PRO	-	expression tag	UNP Q5MYC0
A	1122	GLN	-	expression tag	UNP Q5MYC0
A	1123	PHE	-	expression tag	UNP Q5MYC0
A	1124	GLU	-	expression tag	UNP Q5MYC0
A	1125	LYS	-	expression tag	UNP Q5MYC0
A	1126	GLY	-	expression tag	UNP Q5MYC0
A	1127	GLY	-	expression tag	UNP Q5MYC0
A	1128	GLY	-	expression tag	UNP Q5MYC0
A	1129	SER	-	expression tag	UNP Q5MYC0
A	1130	GLY	-	expression tag	UNP Q5MYC0
A	1131	GLY	-	expression tag	UNP Q5MYC0
A	1132	GLY	-	expression tag	UNP Q5MYC0
A	1133	SER	-	expression tag	UNP Q5MYC0
A	1134	GLY	-	expression tag	UNP Q5MYC0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1135	GLY	-	expression tag	UNP Q5MYC0
A	1136	GLY	-	expression tag	UNP Q5MYC0
A	1137	SER	-	expression tag	UNP Q5MYC0
A	1138	TRP	-	expression tag	UNP Q5MYC0
A	1139	SER	-	expression tag	UNP Q5MYC0
A	1140	HIS	-	expression tag	UNP Q5MYC0
A	1141	PRO	-	expression tag	UNP Q5MYC0
A	1142	GLN	-	expression tag	UNP Q5MYC0
A	1143	PHE	-	expression tag	UNP Q5MYC0
A	1144	GLU	-	expression tag	UNP Q5MYC0
A	1145	LYS	-	expression tag	UNP Q5MYC0

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-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 CADMIUM ION (three-letter code: CD) (formula: Cd) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cd		
3	A	4	4	4	0	0

- Molecule 4 is water.

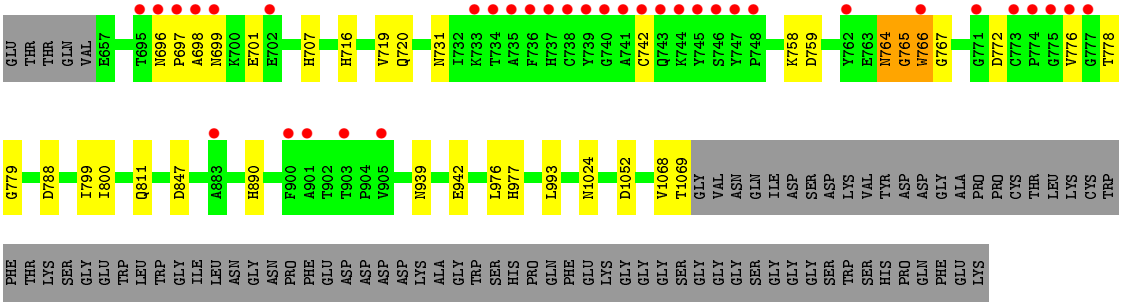
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	83	83	83	0	0

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-(1-3)-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	90.35Å 90.35Å 506.38Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.83 – 2.40 39.83 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (39.83-2.40) 100.0 (39.83-2.40)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.25 (at 2.39Å)	Xtriage
Refinement program	PHENIX 1.14rc3	Depositor
R, R_{free}	0.196 , 0.236 0.200 , 0.240	Depositor DCC
R_{free} test set	1541 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å^2)	55.0	Xtriage
Anisotropy	0.268	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 47.6	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	0.95	EDS
Total number of atoms	3291	wwPDB-VP
Average B, all atoms (Å^2)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.79% 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: CD, 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	0.40	0/3229	0.59	0/4380

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	3154	0	3022	21	0
2	B	50	0	43	0	0
3	A	4	0	0	0	0
4	A	83	0	0	4	0
All	All	3291	0	3065	21	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 3.

The worst 5 of 21 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:701:GLU:OE1	4:A:1301:HOH:O	2.06	0.73
1:A:707:HIS:HB2	1:A:811:GLN:HB3	1.71	0.70
1:A:696:ASN:HB3	1:A:697:PRO:HA	1.75	0.68
1:A:1068:VAL:HG13	1:A:1069:THR:HG23	1.77	0.66
1:A:765:GLY:HA2	1:A:776:VAL:HG23	1.80	0.63

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	411/494 (83%)	392 (95%)	17 (4%)	2 (0%)	29 41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	765	GLY
1	A	766	TRP

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	354/418 (85%)	347 (98%)	7 (2%)	55 74

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	976	LEU
1	A	1024	ASN
1	A	977	HIS
1	A	764	ASN
1	A	993	LEU

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)

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	0.57	0	17,19,21	0.82	1 (5%)
2	NAG	B	2	2	14,14,15	0.62	0	17,19,21	0.70	0
2	BMA	B	3	2	11,11,12	1.14	1 (9%)	15,15,17	1.09	2 (13%)
2	MAN	B	4	2	11,11,12	1.20	1 (9%)	15,15,17	1.20	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	NAG	B	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	2	2	-	2/6/23/26	0/1/1/1
2	BMA	B	3	2	-	2/2/19/22	0/1/1/1
2	MAN	B	4	2	-	2/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	4	MAN	C1-C2	2.58	1.58	1.52
2	B	3	BMA	C2-C3	2.20	1.55	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	4	MAN	C1-O5-C5	2.93	116.16	112.19
2	B	1	NAG	C1-O5-C5	2.63	115.76	112.19
2	B	3	BMA	C1-C2-C3	2.18	112.35	109.67
2	B	3	BMA	C2-C3-C4	2.09	114.51	110.89

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

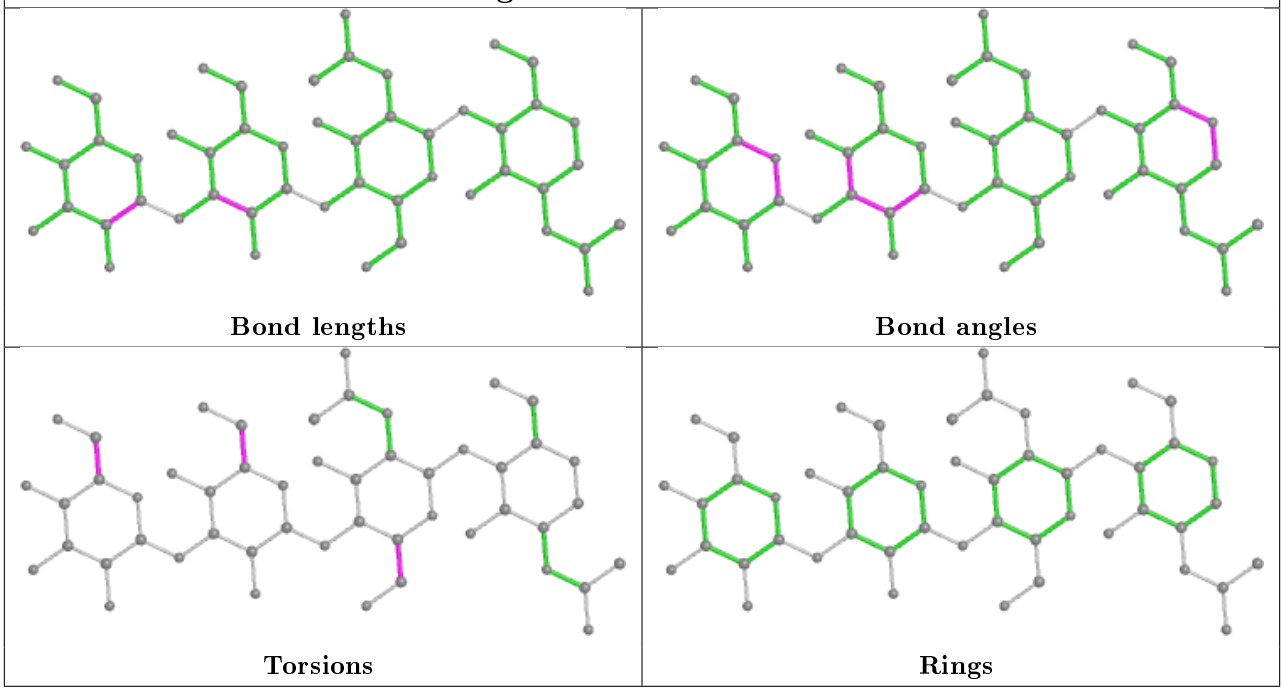
Mol	Chain	Res	Type	Atoms
2	B	3	BMA	O5-C5-C6-O6
2	B	4	MAN	C4-C5-C6-O6
2	B	3	BMA	C4-C5-C6-O6
2	B	4	MAN	O5-C5-C6-O6
2	B	2	NAG	C4-C5-C6-O6

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)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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	413/494 (83%)	0.23	35 (8%)	10	10	40, 57, 144, 208	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	766	TRP	7.4
1	A	900	PHE	7.2
1	A	774	PRO	7.1
1	A	739	TYR	6.5
1	A	744	LYS	6.4

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)

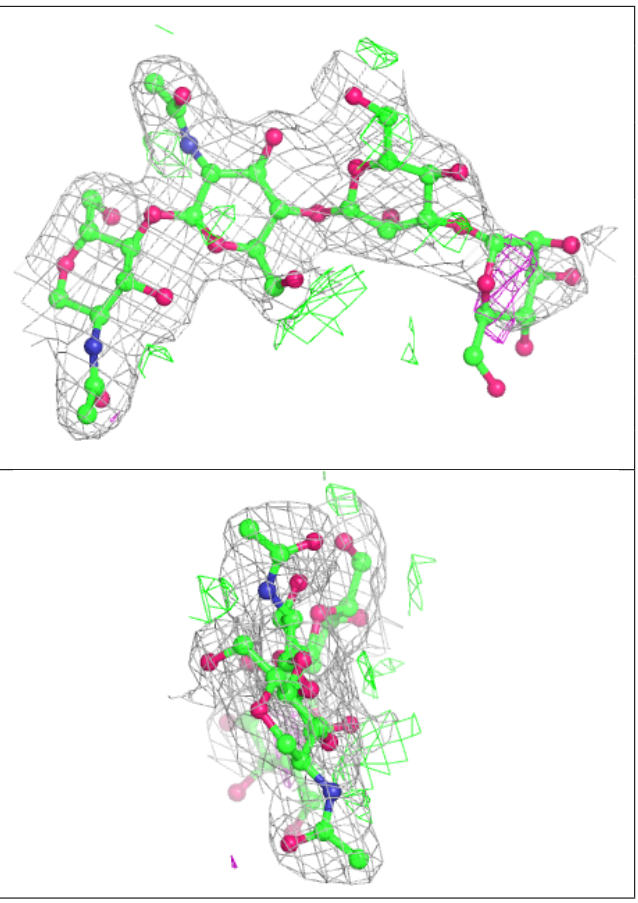
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	MAN	B	4	11/12	0.77	0.61	99,115,122,123	0
2	BMA	B	3	11/12	0.77	0.23	89,100,110,114	0
2	NAG	B	2	14/15	0.89	0.16	45,74,87,97	0
2	NAG	B	1	14/15	0.97	0.09	35,49,66,66	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)

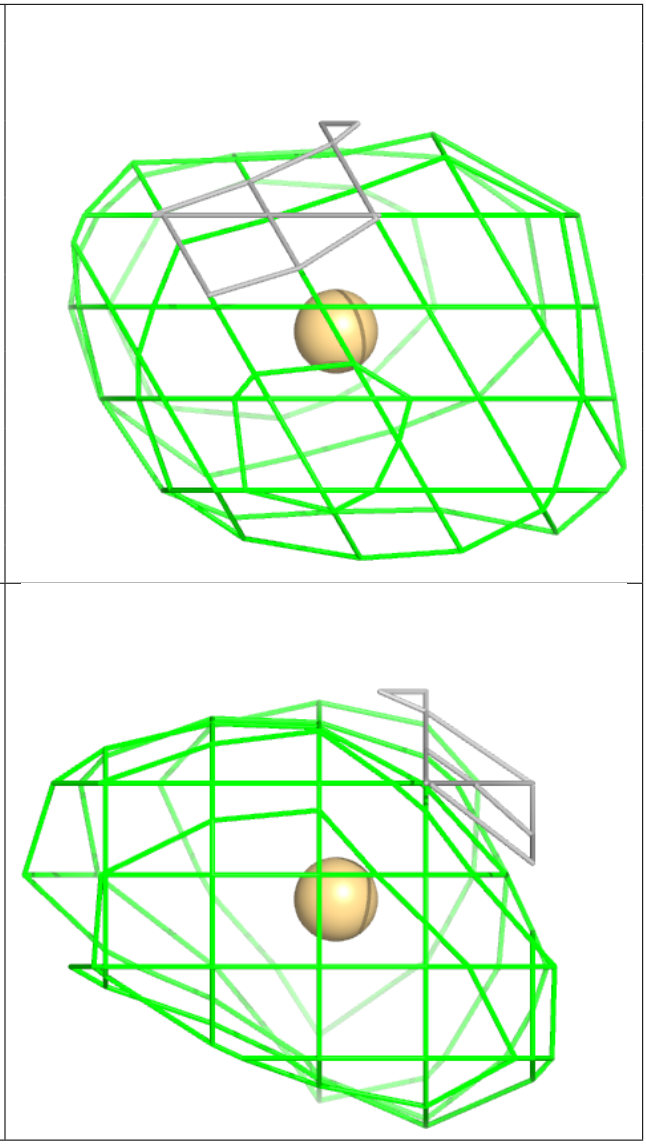
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	CD	A	1208	1/1	0.67	0.68	288,288,288,288	0
3	CD	A	1206	1/1	0.99	0.05	99,99,99,99	0
3	CD	A	1205	1/1	1.00	0.17	58,58,58,58	0
3	CD	A	1207	1/1	1.00	0.18	52,52,52,52	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

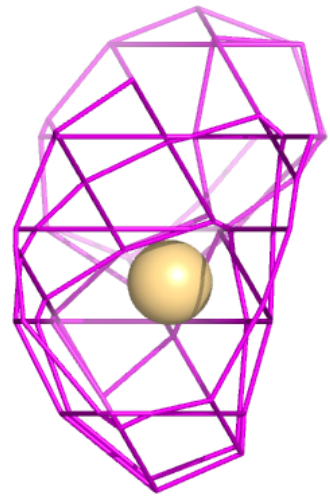
Electron density around CD A 1208:

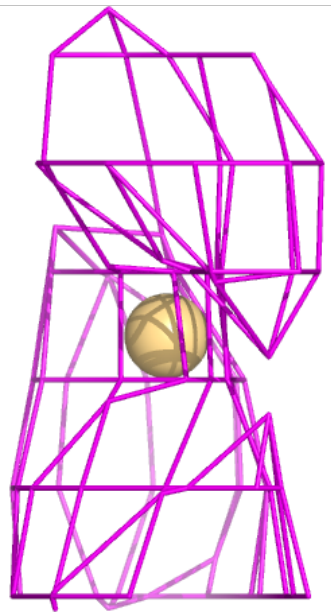
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

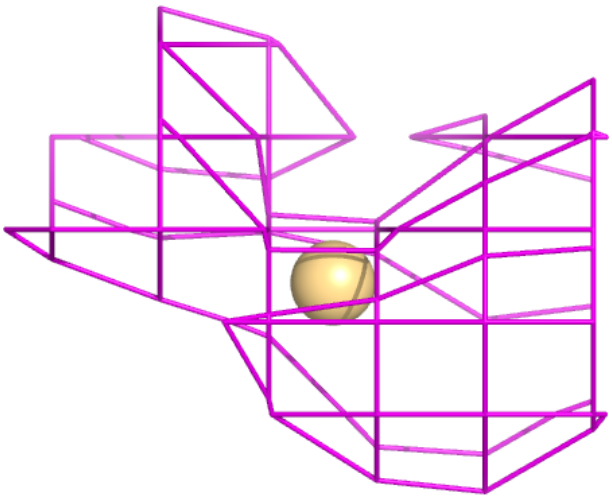


Electron density around CD A 1206:

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

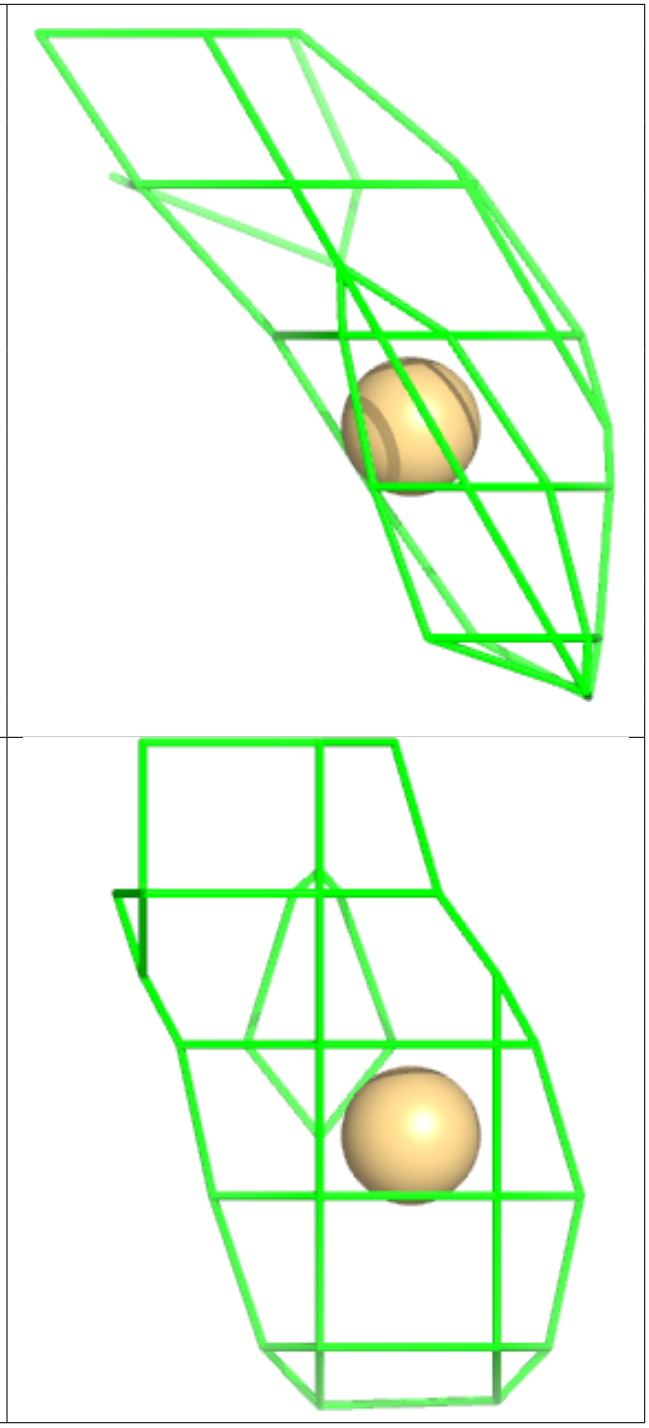






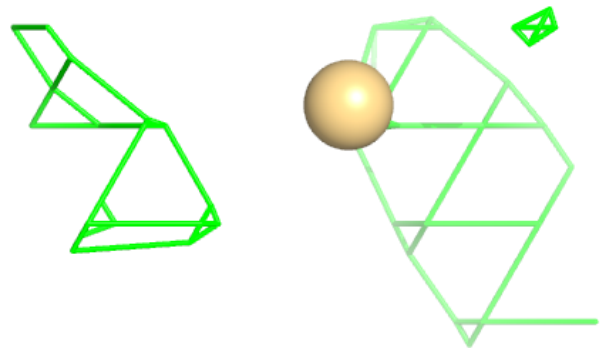
Electron density around CD A 1205:

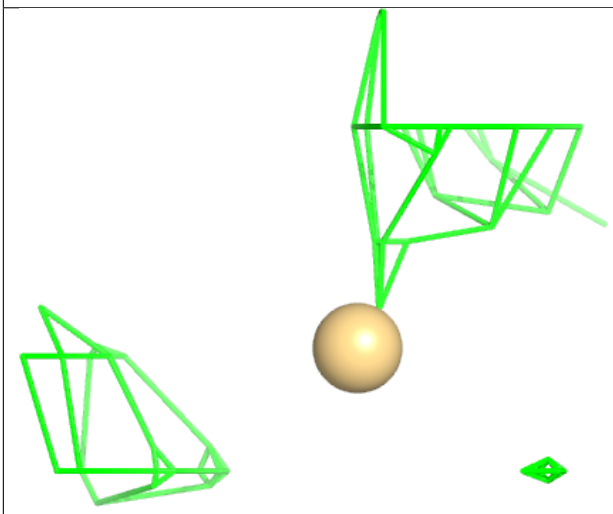
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

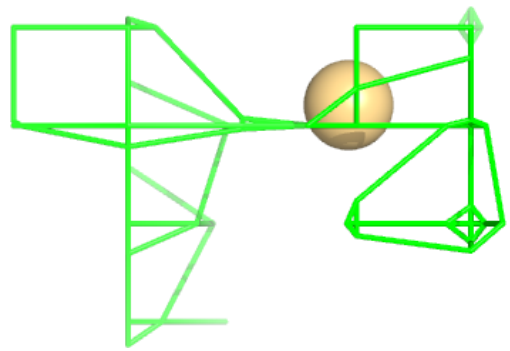


Electron density around CD A 1207:

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







6.5 Other polymers (i)

There are no such residues in this entry.