Full wwPDB X-ray Structure Validation Report ⓘ

Nov 2, 2026 – 03:24 pm GMT

PDB ID	:	29BS / pdb_000029bs
Title	:	Crystal structure of the Norovirus GII.17 P domain in complex with Lacto-N-fucopentaose II (LNFP II)
Authors	:	Pisapia, L.; Varrot, A.
Deposited on	:	2026-03-05
Resolution	:	1.53 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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 ⓘ) were used in the production of this report:

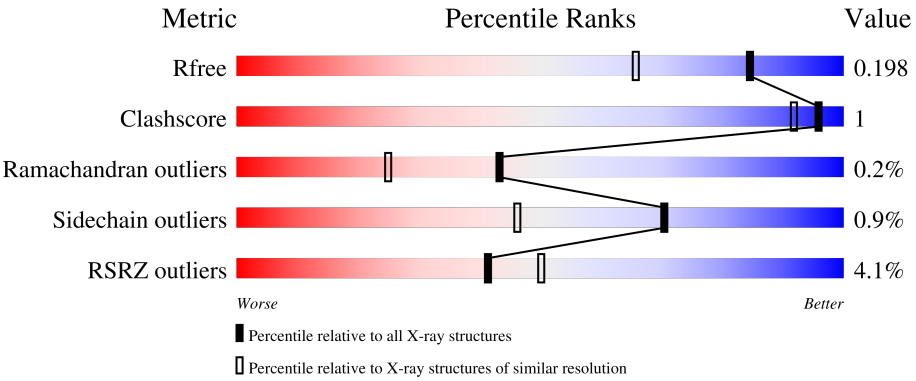
MolProbity	:	4-5-2 with Phenix2.2
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.2
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.51

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 1.53 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}	180053	1003 (1.54-1.54)
Clashscore	190562	1025 (1.54-1.54)
Ramachandran outliers	187476	1007 (1.54-1.54)
Sidechain outliers	187428	1007 (1.54-1.54)
RSRZ outliers	180081	1002 (1.54-1.54)

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	308	5%, 94%, 6%
1	B	308	3%, 94%, 5% ••
2	C	5	40%, 60%
3	D	3	33%, 67%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5704 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.

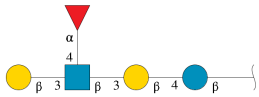
- Molecule 1 is a protein called VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	308	Total	C	N	O	S	0	2	0
			2419	1531	412	466	10			
1	B	305	Total	C	N	O	S	0	0	0
			2385	1512	407	457	9			

There are 24 discrepancies between the modelled and reference sequences:

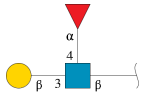
Chain	Residue	Modelled	Actual	Comment	Reference
A	221	GLY	-	expression tag	UNP A0A1C9M2L7
A	222	ALA	-	expression tag	UNP A0A1C9M2L7
A	223	MET	-	expression tag	UNP A0A1C9M2L7
A	224	ALA	-	expression tag	UNP A0A1C9M2L7
A	334	MET	VAL	conflict	UNP A0A1C9M2L7
A	350	SER	ALA	conflict	UNP A0A1C9M2L7
A	361	ARG	LYS	conflict	UNP A0A1C9M2L7
A	372	LYS	ARG	conflict	UNP A0A1C9M2L7
A	396	GLN	HIS	conflict	UNP A0A1C9M2L7
A	431	PHE	LEU	conflict	UNP A0A1C9M2L7
A	435	SER	PHE	conflict	UNP A0A1C9M2L7
A	469	GLU	ASP	conflict	UNP A0A1C9M2L7
B	221	GLY	-	expression tag	UNP A0A1C9M2L7
B	222	ALA	-	expression tag	UNP A0A1C9M2L7
B	223	MET	-	expression tag	UNP A0A1C9M2L7
B	224	ALA	-	expression tag	UNP A0A1C9M2L7
B	334	MET	VAL	conflict	UNP A0A1C9M2L7
B	350	SER	ALA	conflict	UNP A0A1C9M2L7
B	361	ARG	LYS	conflict	UNP A0A1C9M2L7
B	372	LYS	ARG	conflict	UNP A0A1C9M2L7
B	396	GLN	HIS	conflict	UNP A0A1C9M2L7
B	431	PHE	LEU	conflict	UNP A0A1C9M2L7
B	435	SER	PHE	conflict	UNP A0A1C9M2L7
B	469	GLU	ASP	conflict	UNP A0A1C9M2L7

- Molecule 2 is an oligosaccharide called beta-D-galactopyranose-(1-3)-[alpha-L-fucopyranose-(1-4)]2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	5	Total	C	N	O	0	0	0
			58	32	1	25			

- Molecule 3 is an oligosaccharide called beta-D-galactopyranose-(1-3)-[alpha-L-fucopyranose-(1-4)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	3	Total	C	N	O	0	0	0
			36	20	1	15			

- Molecule 4 is water.

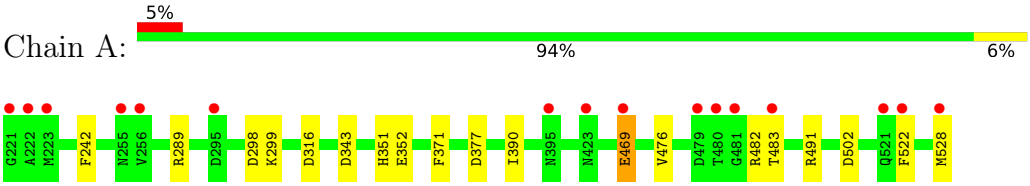
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	413	Total	O	0	1
			414	414		
4	B	391	Total	O	0	1
			392	392		

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.

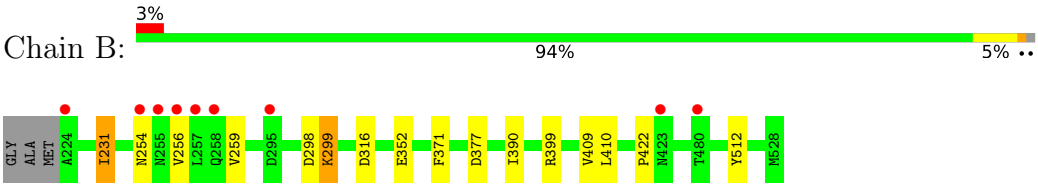
• Molecule 1: VP1

Chain A:



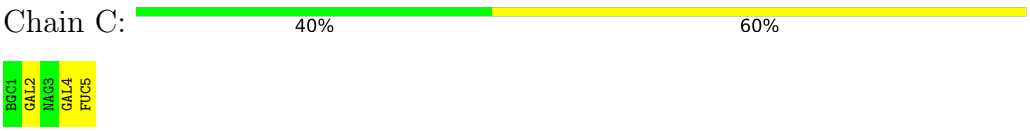
• Molecule 1: VP1

Chain B:



• Molecule 2: beta-D-galactopyranose-(1-3)-[alpha-L-fucopyranose-(1-4)]2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain C:



• Molecule 3: beta-D-galactopyranose-(1-3)-[alpha-L-fucopyranose-(1-4)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	59.50Å 86.27Å 63.97Å 90.00° 106.08° 90.00°	Depositor
Resolution (Å)	42.78 – 1.53 42.78 – 1.53	Depositor EDS
% Data completeness (in resolution range)	99.8 (42.78-1.53) 99.8 (42.78-1.53)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 1.53Å)	Xtriage
Refinement program	REFMAC 5.8.0430	Depositor
R, R_{free}	0.159 , 0.188 0.173 , 0.198	Depositor DCC
R_{free} test set	4687 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	16.7	Xtriage
Anisotropy	0.390	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 34.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	5704	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 4.75% 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: BGC, GAL, NAG, FUC

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	0.76	1/2488 (0.0%)	1.13	9/3399 (0.3%)
1	B	0.77	0/2454	1.13	6/3354 (0.2%)
All	All	0.76	1/4942 (0.0%)	1.13	15/6753 (0.2%)

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	A	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	351	HIS	CE1-NE2	-5.58	1.26	1.32

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	377	ASP	CA-CB-CG	8.26	120.86	112.60
1	A	469	GLU	CB-CG-CD	7.50	125.36	112.60
1	A	343	ASP	CA-CB-CG	7.42	120.02	112.60
1	A	298	ASP	CA-CB-CG	7.18	119.78	112.60
1	B	316	ASP	CA-CB-CG	6.46	119.06	112.60
1	A	502	ASP	CA-CB-CG	6.35	118.95	112.60
1	A	371	PHE	CA-CB-CG	5.68	119.48	113.80
1	B	298	ASP	CA-CB-CG	5.55	118.16	112.60
1	A	242	PHE	CA-CB-CG	-5.54	108.26	113.80
1	B	371	PHE	CA-CB-CG	5.35	119.15	113.80

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	522	PHE	CA-CB-CG	-5.11	108.69	113.80
1	B	399	ARG	NE-CZ-NH1	-5.09	116.41	121.50
1	B	299	LYS	N-CA-CB	-5.08	102.64	110.16
1	A	377	ASP	CA-CB-CG	5.07	117.67	112.60
1	A	316	ASP	CA-CB-CG	5.03	117.63	112.60

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	289	ARG	Sidechain
1	A	482	ARG	Sidechain
1	A	491	ARG	Sidechain

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	2419	0	2304	6	0
1	B	2385	0	2279	7	0
2	C	58	0	51	0	0
3	D	36	0	33	0	0
4	A	414	0	0	0	0
4	B	392	0	0	0	0
All	All	5704	0	4667	10	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 1.

All (10) 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:352[A]:GLU:HG3	1:B:352:GLU:HG3	1.76	0.66
1:B:409:VAL:HG12	1:B:410:LEU:HG	1.88	0.55
1:A:390:ILE:C	1:A:390:ILE:HD12	2.37	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:390:ILE:C	1:B:390:ILE:HD12	2.39	0.48
1:A:476:VAL:HG12	1:A:483:THR:HG22	1.95	0.48
1:A:352[A]:GLU:HG3	1:B:352:GLU:CG	2.44	0.45
1:B:254:ASN:OD1	1:B:422:PRO:HG2	2.20	0.42
1:A:528:MET:HE2	1:A:528:MET:HB3	1.90	0.41
1:A:352[A]:GLU:CG	1:B:352:GLU:HG3	2.48	0.41
1:B:231:ILE:HG23	1:B:512:TYR:CZ	2.55	0.41

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	308/308 (100%)	302 (98%)	6 (2%)	0	100 100
1	B	303/308 (98%)	297 (98%)	5 (2%)	1 (0%)	36 19
All	All	611/616 (99%)	599 (98%)	11 (2%)	1 (0%)	43 25

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	259	VAL

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	269/267 (101%)	267 (99%)	2 (1%)	76 55
1	B	266/267 (100%)	263 (99%)	3 (1%)	65 38
All	All	535/534 (100%)	530 (99%)	5 (1%)	70 46

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

Mol	Chain	Res	Type
1	A	299	LYS
1	A	469	GLU
1	B	231	ILE
1	B	256	VAL
1	B	299	LYS

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	296	ASN

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)

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	BGC	C	1	2	12,12,12	0.34	0	17,17,17	0.67	0
2	GAL	C	2	2	11,11,12	0.70	0	15,15,17	1.05	2 (13%)
2	NAG	C	3	2	14,14,15	0.35	0	17,19,21	0.82	0
2	GAL	C	4	2	11,11,12	0.92	1 (9%)	15,15,17	1.07	1 (6%)
2	FUC	C	5	2	10,10,11	0.90	0	14,14,16	1.14	1 (7%)
3	NAG	D	1	3	15,15,15	0.47	0	21,21,21	0.84	0
3	GAL	D	2	3	11,11,12	0.35	0	15,15,17	1.09	1 (6%)
3	FUC	D	3	3	10,10,11	1.20	1 (10%)	14,14,16	0.95	1 (7%)

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	BGC	C	1	2	-	0/2/22/22	0/1/1/1
2	GAL	C	2	2	-	0/2/19/22	0/1/1/1
2	NAG	C	3	2	-	0/6/23/26	0/1/1/1
2	GAL	C	4	2	-	0/2/19/22	0/1/1/1
2	FUC	C	5	2	-	-	0/1/1/1
3	NAG	D	1	3	-	2/6/26/26	0/1/1/1
3	GAL	D	2	3	-	0/2/19/22	0/1/1/1
3	FUC	D	3	3	-	-	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	3	FUC	C2-C3	3.49	1.57	1.52
2	C	4	GAL	O5-C5	2.68	1.48	1.43

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	2	GAL	C1-C2-C3	3.39	113.83	109.67
2	C	4	GAL	C1-C2-C3	3.12	113.50	109.67
2	C	2	GAL	C1-C2-C3	2.32	112.51	109.67
2	C	2	GAL	O3-C3-C2	-2.16	105.86	109.99
2	C	5	FUC	O2-C2-C3	-2.16	105.82	110.14
3	D	3	FUC	O2-C2-C3	-2.07	105.99	110.14

There are no chirality outliers.

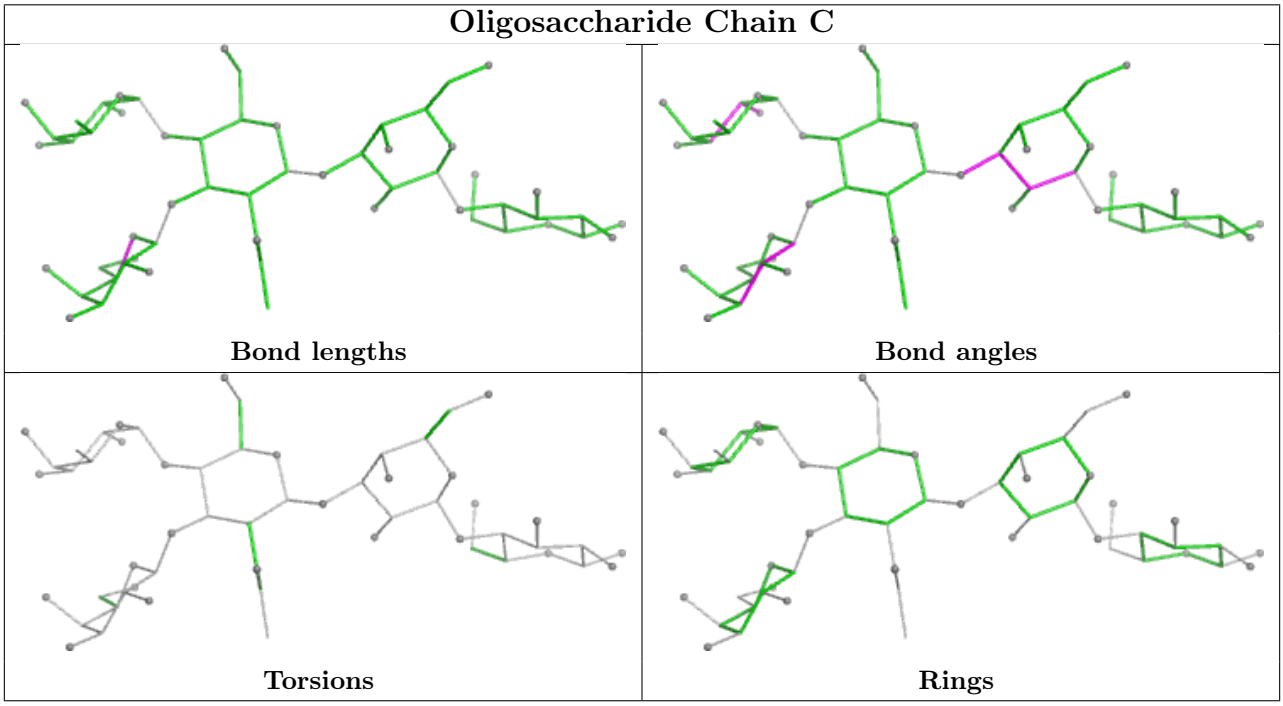
All (2) torsion outliers are listed below:

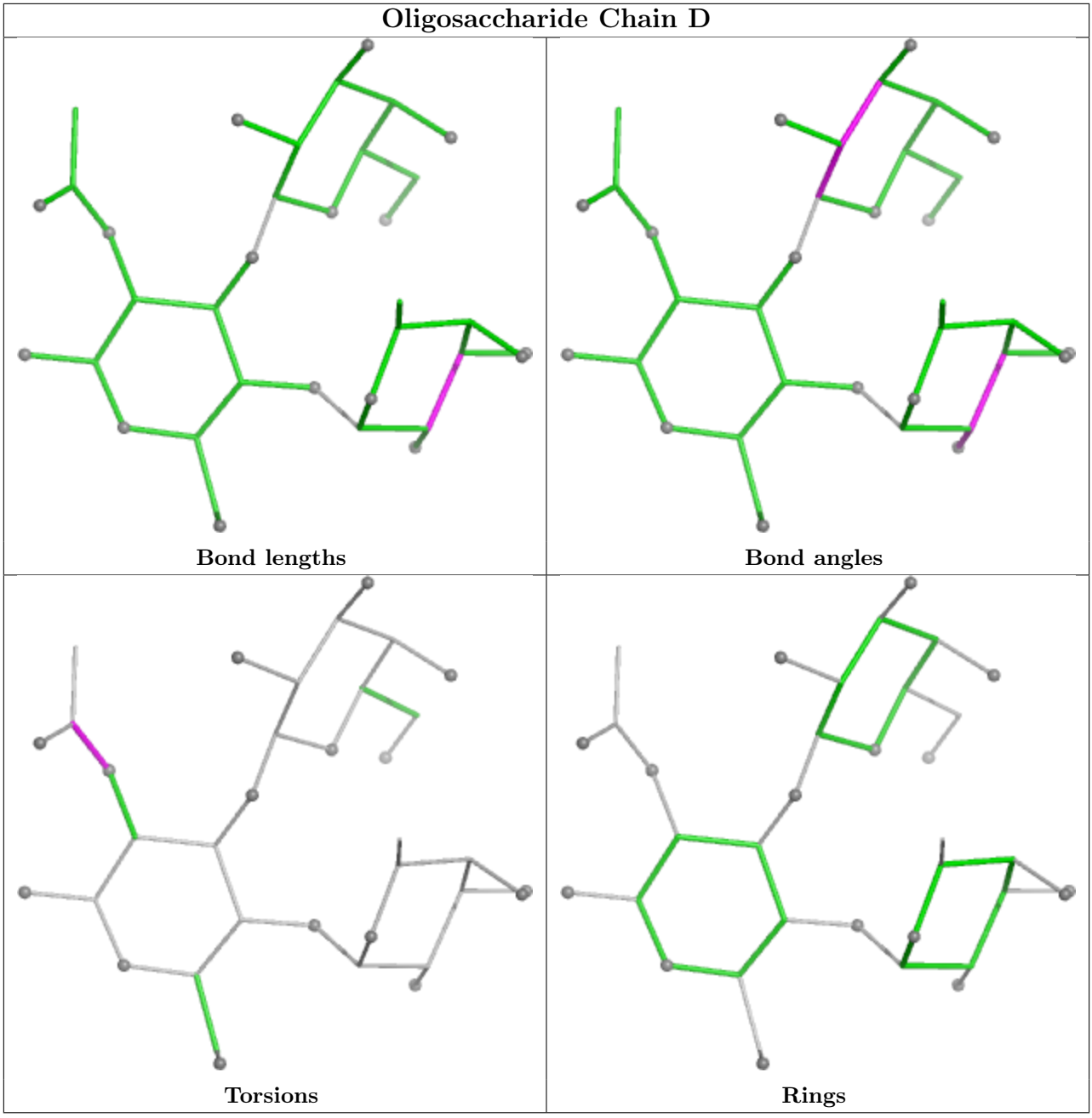
Mol	Chain	Res	Type	Atoms
3	D	1	NAG	C8-C7-N2-C2
3	D	1	NAG	O7-C7-N2-C2

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.





5.6 Ligand geometry (i)

There are no ligands in this entry.

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	308/308 (100%)	0.05	16 (5%) 33 40	7, 15, 36, 55	2 (0%)
1	B	305/308 (99%)	-0.03	9 (2%) 52 61	10, 16, 30, 57	0
All	All	613/616 (99%)	0.01	25 (4%) 41 50	7, 15, 36, 57	2 (0%)

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	256	VAL	5.4
1	A	221	GLY	4.5
1	B	257	LEU	4.3
1	B	423	ASN	4.2
1	B	295	ASP	4.0
1	A	522	PHE	3.8
1	A	256	VAL	3.6
1	B	224	ALA	3.5
1	A	223	MET	3.5
1	B	480	THR	3.3
1	A	480	THR	3.2
1	A	222	ALA	3.1
1	B	258	GLN	3.0
1	A	479	ASP	2.9
1	B	255	ASN	2.8
1	A	255	ASN	2.7
1	A	521	GLN	2.5
1	A	423	ASN	2.3
1	A	295	ASP	2.2
1	A	469	GLU	2.1
1	A	481	GLY	2.1
1	A	528	MET	2.1
1	B	254	ASN	2.0
1	A	483	THR	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	395	ASN	2.0

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)

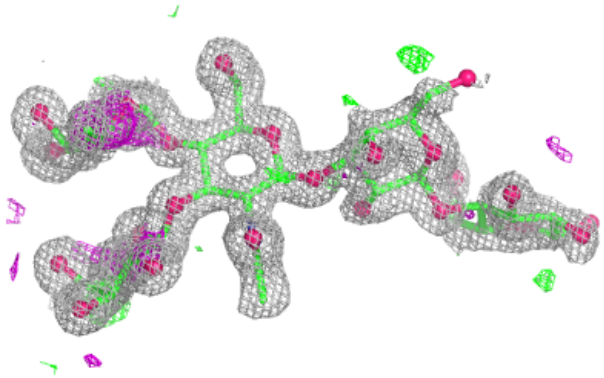
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

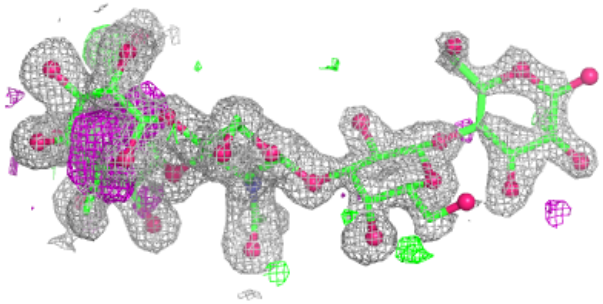
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	GAL	D	2	11/12	0.67	0.25	25,29,30,32	11
2	BGC	C	1	12/12	0.76	0.14	24,31,36,37	12
3	NAG	D	1	15/15	0.84	0.15	25,30,40,47	0
2	GAL	C	2	11/12	0.84	0.13	18,21,24,32	11
2	GAL	C	4	11/12	0.90	0.10	22,23,30,34	0
2	FUC	C	5	10/11	0.90	0.11	18,18,20,20	0
3	FUC	D	3	10/11	0.92	0.10	18,22,25,25	0
2	NAG	C	3	14/15	0.94	0.08	19,21,23,25	0

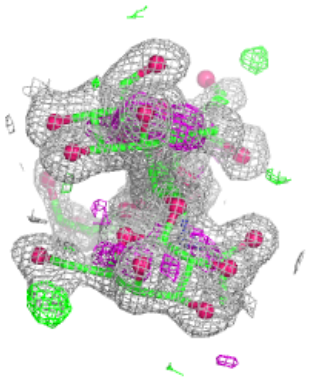
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

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)

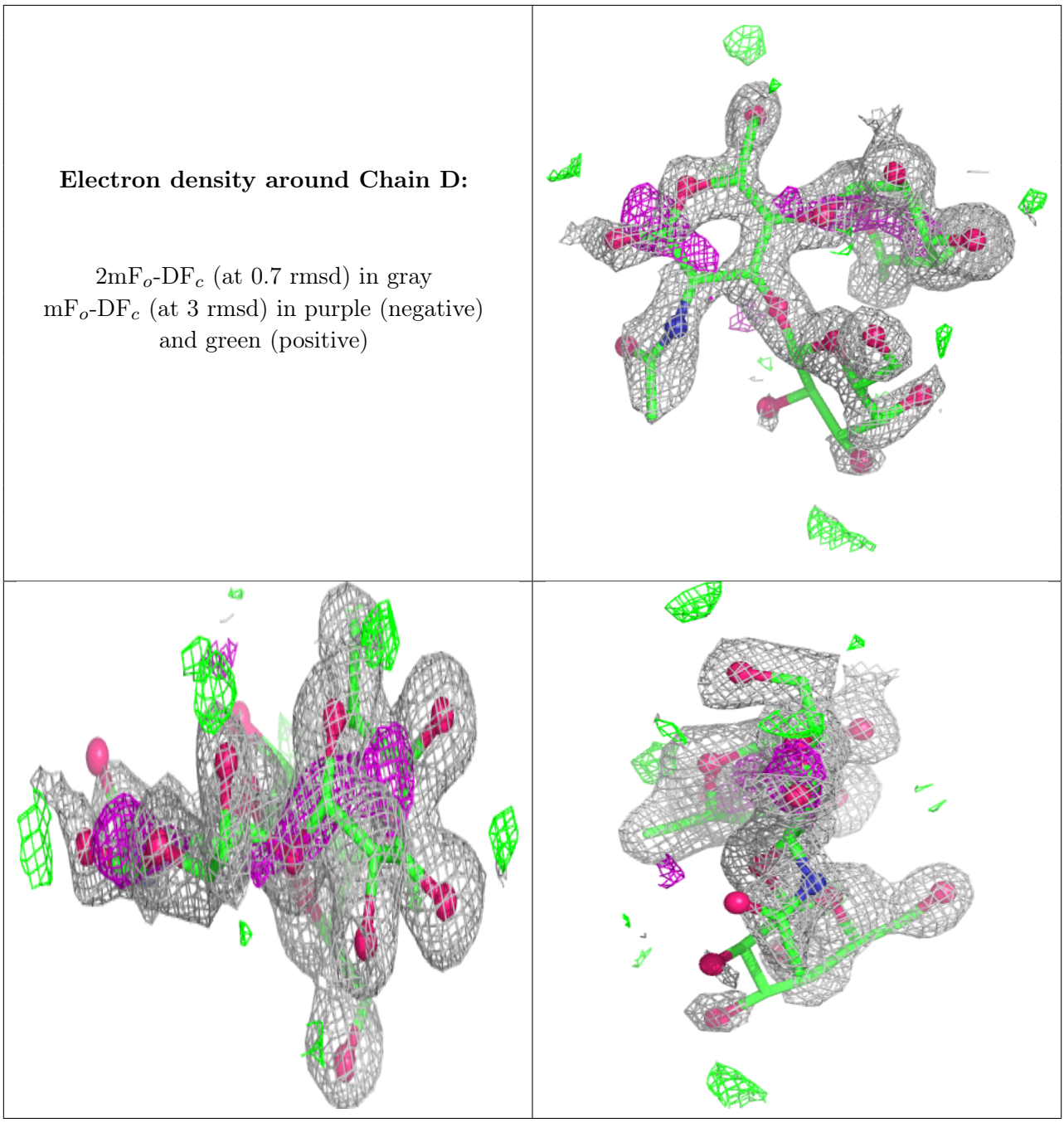






Electron density around Chain D:

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



6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.