



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 31, 2022 – 12:09 PM JST

PDB ID : 7VS8  
Title : Crystal structure of P domain from norovirus GI.9 capsid protein in complex with Lewis b antigen.  
Authors : Murayama, K.; Kato-Murayama, M.; Shirouzu, M.  
Deposited on : 2021-10-26  
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](mailto: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	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.30
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)	:	2.30

i

## X-RAY DIFFRACTION

A.

the following graphic. The table shows the number of entries on which the scores are based.

R. f. m. c.

the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled.

1

## 2 Entry composition [i](#)

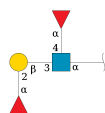
There are 5 unique types of molecules in this entry. The entry contains 5217 atoms, of which 0 are hydrogens and 0 are deuteriums.

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	310	Total	C	N	O	S	0	0	0
			2369	1517	401	444	7			
1	B	310	Total	C	N	O	S	0	0	0
			2369	1517	401	444	7			

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	4	Total	C	N	O	0	0	0
			46	26	1	19			
2	D	4	Total	C	N	O	0	0	0
			46	26	1	19			

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	5	Total	Cl	0	0
			5	5		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	6	Total	Cl	0	0
			6	6		

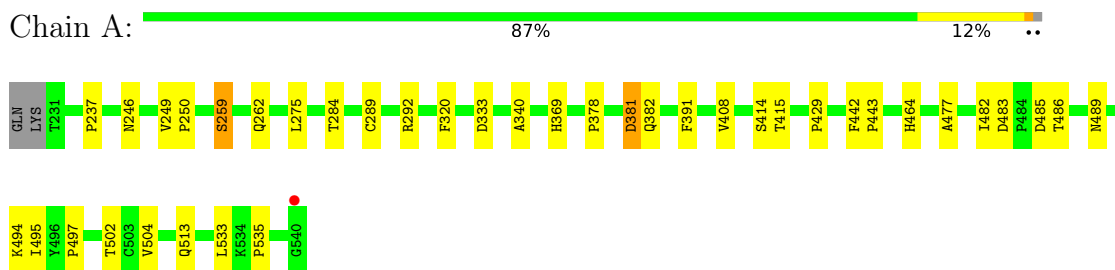
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	187	Total	O	0	0
			187	187		
5	B	187	Total	O	0	0
			187	187		

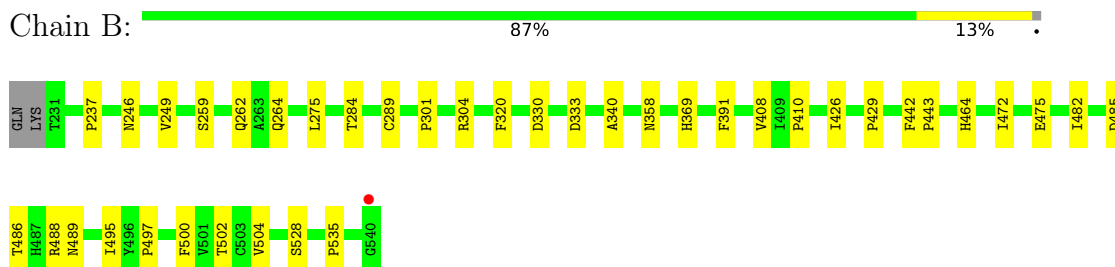
### 3 Residue-property plots

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



- Molecule 1: VP1



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



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



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	183.75Å 76.46Å 103.28Å 90.00° 124.20° 90.00°	Depositor
Resolution (Å)	41.19 – 2.40 41.19 – 2.40	Depositor EDS
% Data completeness (in resolution range)	98.1 (41.19-2.40) 98.1 (41.19-2.40)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.47 (at 2.39Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, $R_{free}$	0.165 , 0.196 0.165 , 0.196	Depositor DCC
$R_{free}$ test set	1995 reflections (4.37%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.9	Xtriage
Anisotropy	0.797	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 33.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.487 for -h-2*k,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5217	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.66% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: CL, FUC, GAL, NDG, MG

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	0.43	0/2442	0.60	0/3356
1	B	0.42	0/2442	0.59	0/3356
All	All	0.43	0/4884	0.59	0/6712

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	2369	0	2288	22	0
1	B	2369	0	2287	20	0
2	C	46	0	39	0	0
2	D	46	0	39	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	5	0	0	0	0
4	B	6	0	0	0	0
5	A	187	0	0	1	0
5	B	187	0	0	0	0
All	All	5217	0	4653	42	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 42 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:B:429:PRO:HG3	1:B:502:THR:HG21	1.69	0.74
1:A:485:ASP:OD1	1:A:486:THR:N	2.30	0.62
1:A:320:PHE:H	1:A:369:HIS:CE1	2.22	0.58
1:B:486:THR:OG1	1:B:488:ARG:HG2	2.04	0.57
1:B:284:THR:HG21	1:B:289:CYS:HB2	1.88	0.56

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	308/312 (99%)	296 (96%)	12 (4%)	0	100	100
1	B	308/312 (99%)	297 (96%)	11 (4%)	0	100	100
All	All	616/624 (99%)	593 (96%)	23 (4%)	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	260/262 (99%)	258 (99%)	2 (1%)	81	91
1	B	260/262 (99%)	258 (99%)	2 (1%)	81	91
All	All	520/524 (99%)	516 (99%)	4 (1%)	81	91

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

Mol	Chain	Res	Type
1	A	259	SER
1	A	381	ASP
1	B	259	SER
1	B	358	ASN

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	262	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

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	NDG	C	1	2	15,15,15	0.62	0	21,21,21	0.89	1 (4%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GAL	C	2	2	11,11,12	0.71	0	15,15,17	1.01	1 (6%)
2	FUC	C	3	2	10,10,11	1.47	1 (10%)	14,14,16	1.05	1 (7%)
2	FUC	C	4	2	10,10,11	1.06	0	14,14,16	0.90	0
2	NDG	D	1	2	15,15,15	0.66	0	21,21,21	0.81	1 (4%)
2	GAL	D	2	2	11,11,12	0.77	0	15,15,17	1.06	1 (6%)
2	FUC	D	3	2	10,10,11	1.18	1 (10%)	14,14,16	1.01	1 (7%)
2	FUC	D	4	2	10,10,11	1.12	0	14,14,16	0.74	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	NDG	C	1	2	-	0/6/26/26	0/1/1/1
2	GAL	C	2	2	-	0/2/19/22	0/1/1/1
2	FUC	C	3	2	-	-	0/1/1/1
2	FUC	C	4	2	-	-	0/1/1/1
2	NDG	D	1	2	-	0/6/26/26	0/1/1/1
2	GAL	D	2	2	-	0/2/19/22	0/1/1/1
2	FUC	D	3	2	-	-	0/1/1/1
2	FUC	D	4	2	-	-	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	3	FUC	C2-C3	3.79	1.58	1.52
2	D	3	FUC	C2-C3	3.00	1.56	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1	NDG	O4-C4-C3	-2.46	104.66	110.35
2	D	1	NDG	O4-C4-C3	-2.37	104.86	110.35
2	C	3	FUC	O3-C3-C4	-2.19	105.28	110.35
2	D	2	GAL	C1-O5-C5	2.09	115.02	112.19
2	C	2	GAL	O3-C3-C4	-2.04	105.63	110.35

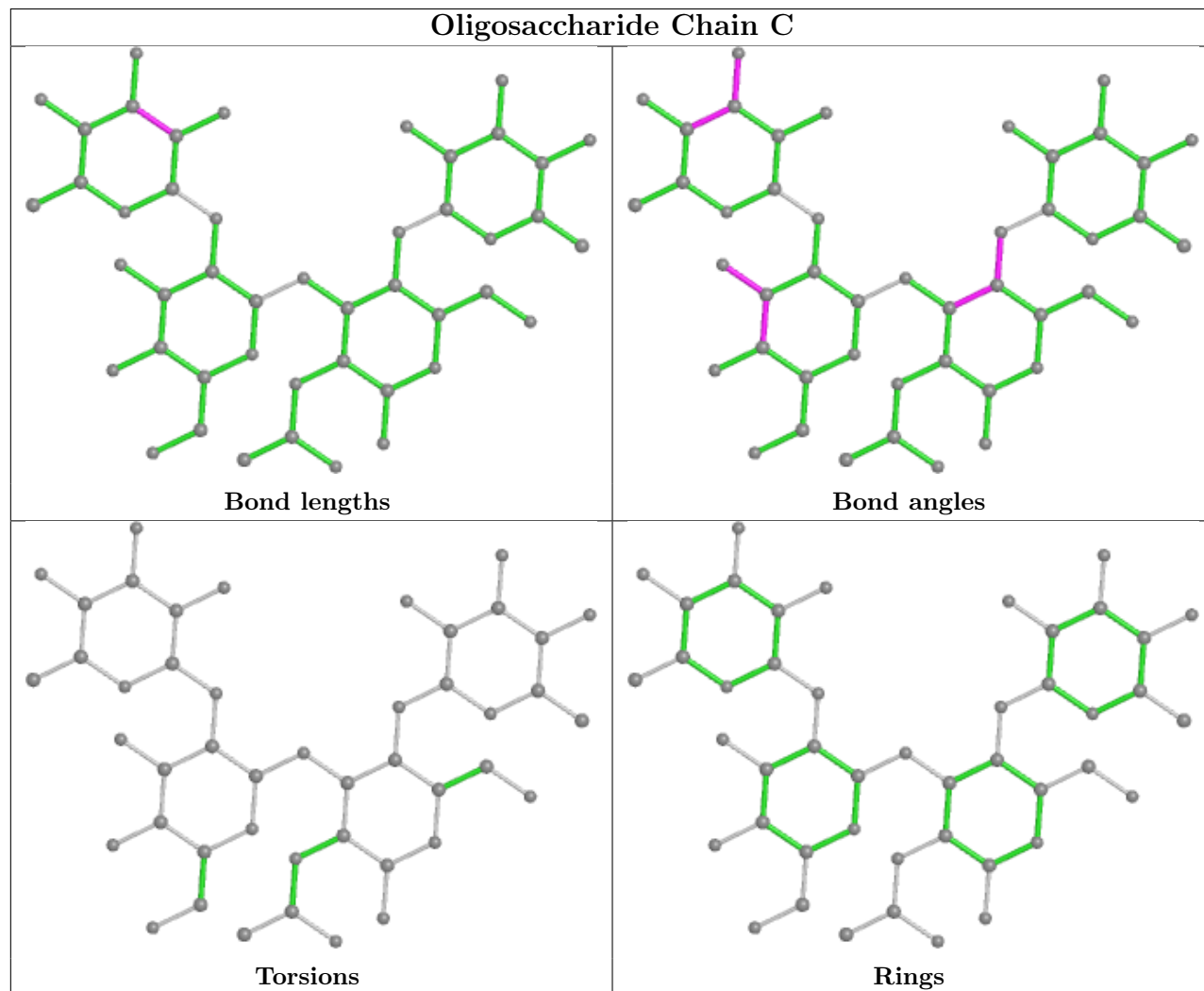
There are no chirality outliers.

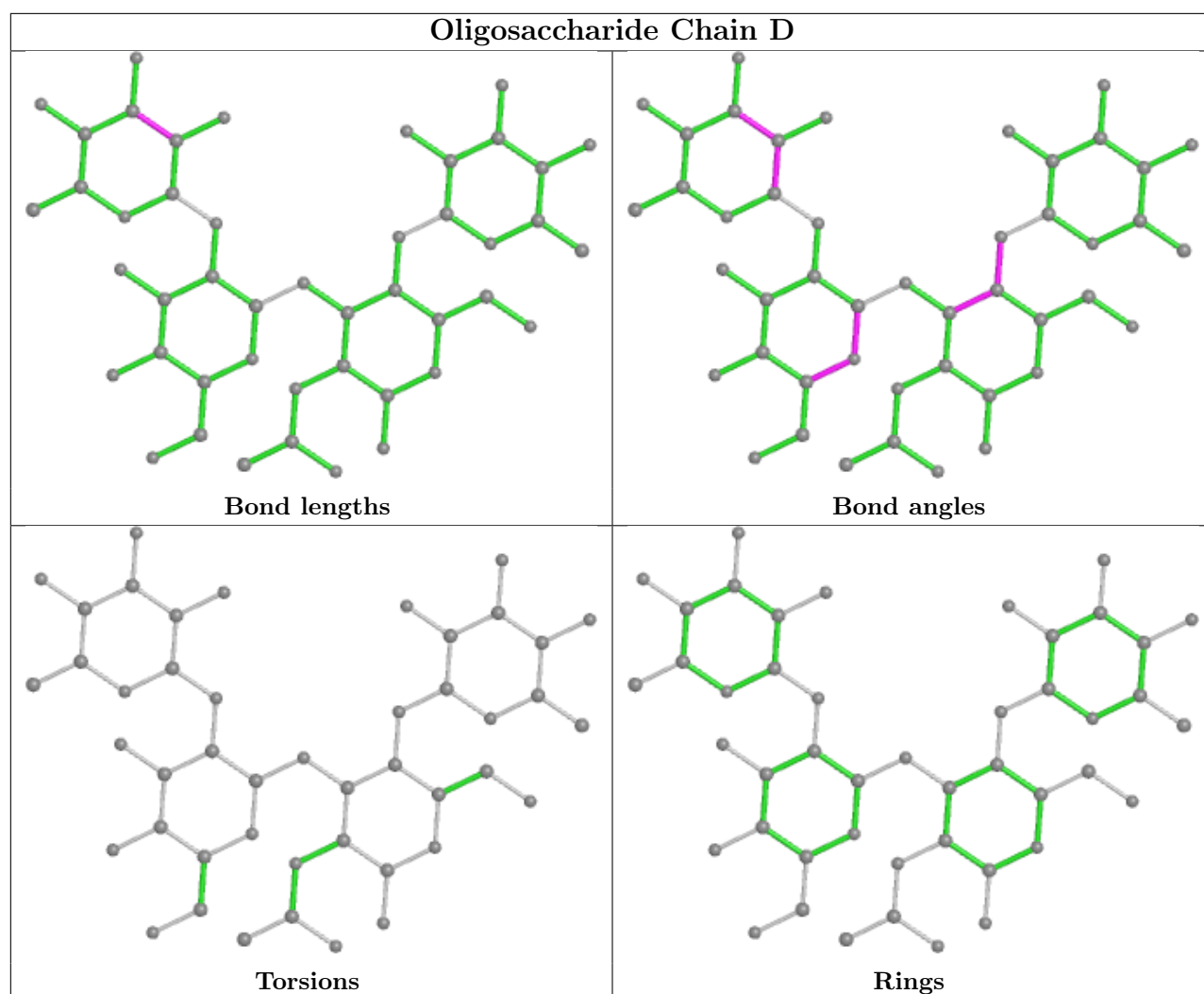
There are no torsion outliers.

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.





## 5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 13 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 ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	310/312 (99%)	-0.37	1 (0%) 94 93	26, 35, 61, 77	0
1	B	310/312 (99%)	-0.36	1 (0%) 94 93	26, 35, 63, 78	0
All	All	620/624 (99%)	-0.37	2 (0%) 94 93	26, 35, 62, 78	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	540	GLY	3.0
1	A	540	GLY	2.8

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

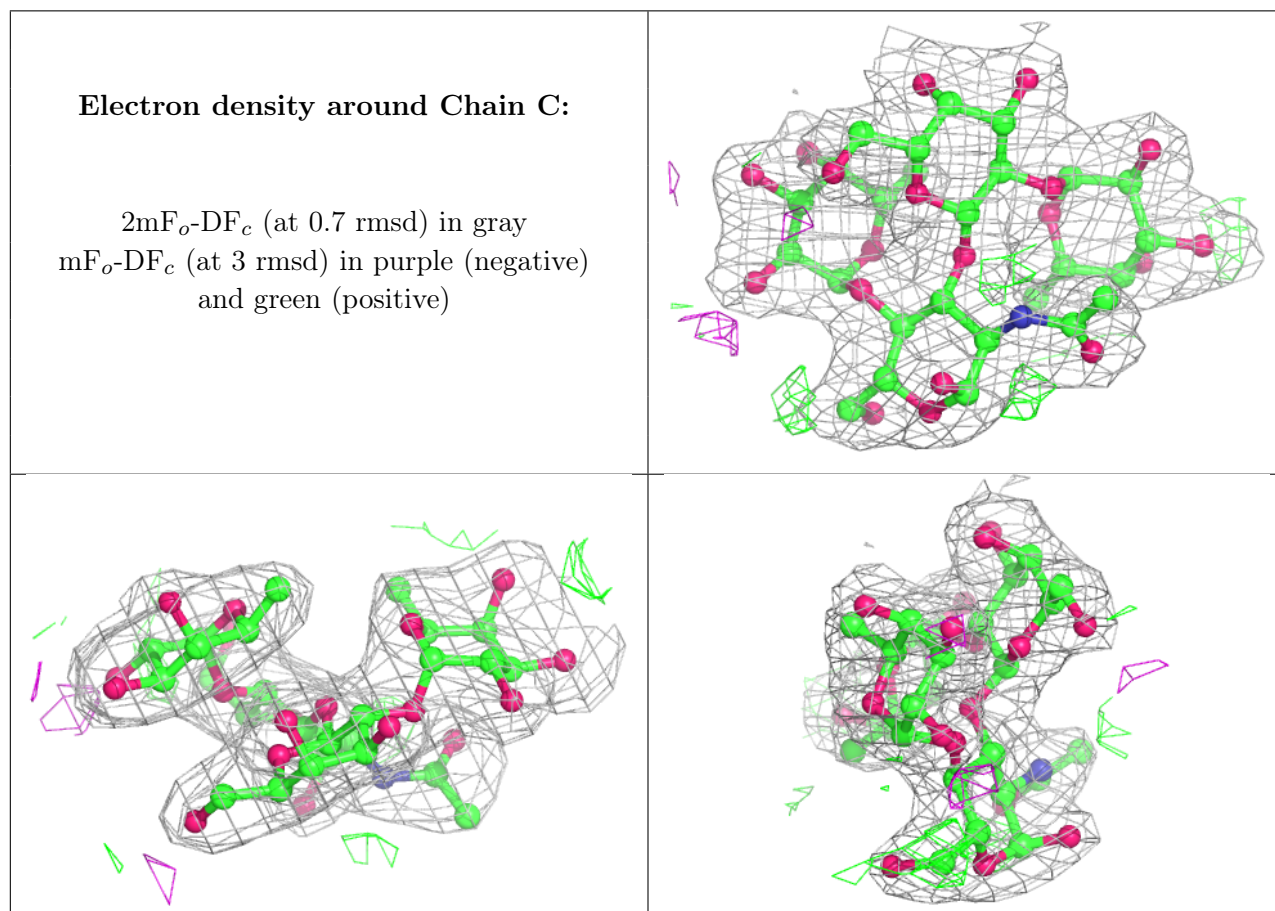
There are no non-standard protein/DNA/RNA residues in this entry.

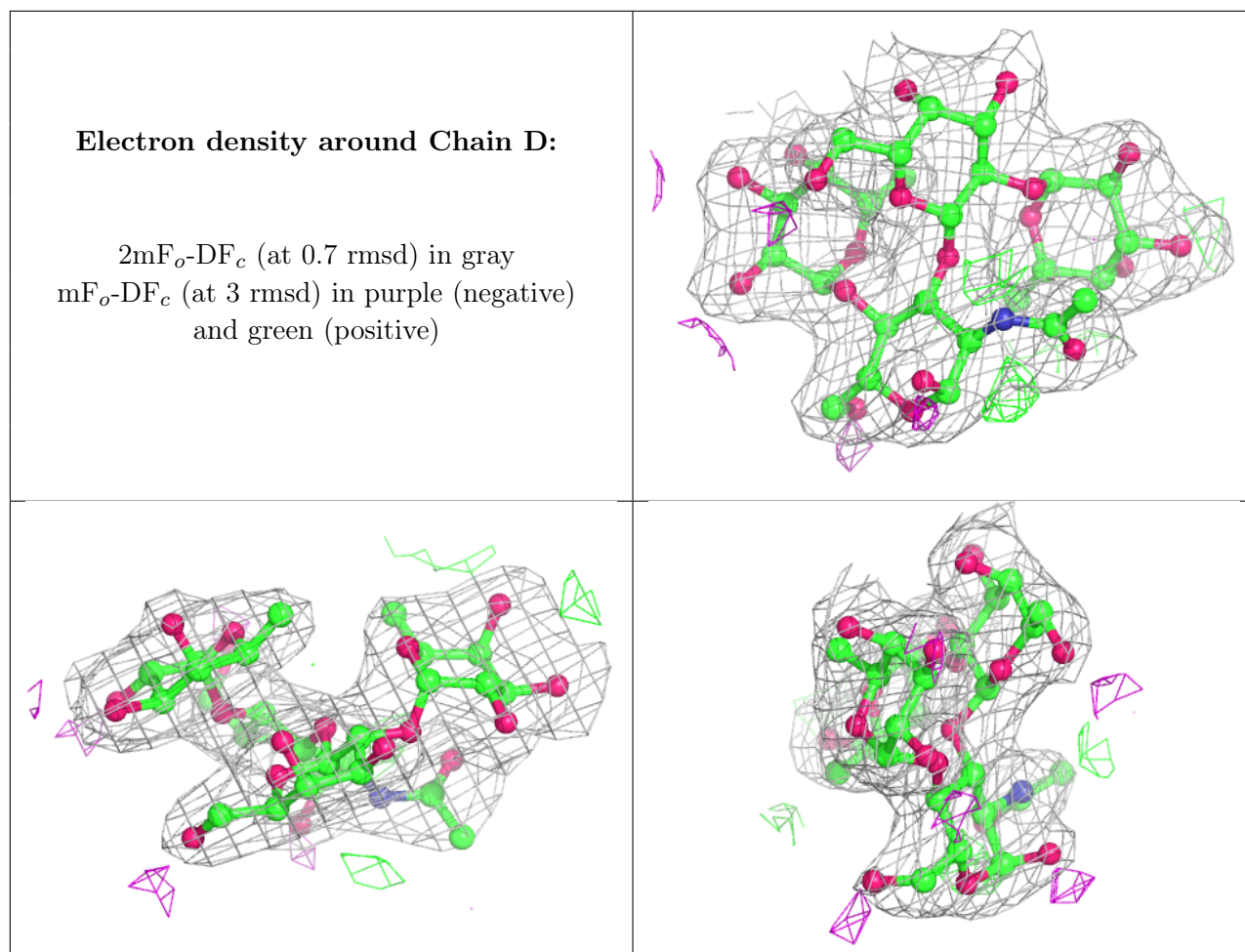
### 6.3 Carbohydrates [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	NDG	C	1	15/15	0.96	0.11	36,41,47,47	0
2	NDG	D	1	15/15	0.96	0.12	39,42,47,49	0
2	FUC	C	3	10/11	0.98	0.11	36,37,41,42	0
2	FUC	C	4	10/11	0.98	0.12	38,41,43,45	0
2	GAL	C	2	11/12	0.98	0.12	35,39,43,46	0
2	FUC	D	3	10/11	0.98	0.12	36,38,40,42	0
2	FUC	D	4	10/11	0.98	0.10	37,41,44,45	0
2	GAL	D	2	11/12	0.99	0.12	34,39,43,48	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.





## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
4	CL	A	604	1/1	0.95	0.07	54,54,54,54	0
4	CL	B	602	1/1	0.96	0.07	50,50,50,50	0
4	CL	B	603	1/1	0.97	0.07	48,48,48,48	0
4	CL	A	605	1/1	0.98	0.06	49,49,49,49	0
4	CL	B	606	1/1	0.98	0.09	48,48,48,48	0
3	MG	B	601	1/1	0.99	0.05	43,43,43,43	0
4	CL	A	606	1/1	0.99	0.09	33,33,33,33	0
4	CL	A	602	1/1	0.99	0.08	56,56,56,56	0
4	CL	A	603	1/1	0.99	0.10	35,35,35,35	0
4	CL	B	605	1/1	0.99	0.08	57,57,57,57	0
3	MG	A	601	1/1	0.99	0.06	42,42,42,42	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	CL	B	607	1/1	0.99	0.10	32,32,32,32	0
4	CL	B	604	1/1	1.00	0.08	36,36,36,36	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.