



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 29, 2021 – 12:35 AM JST

PDB ID : 7CAN
Title : Structure of sybody MR17-K99Y in complex with the SARS-CoV-2 S Receptor-binding domain (RBD)
Authors : Li, T.; Yao, H.; Cai, H.; Qin, W.; Li, D.
Deposited on : 2020-06-09
Resolution : 2.94 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 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 : 2.25
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) : 2.25

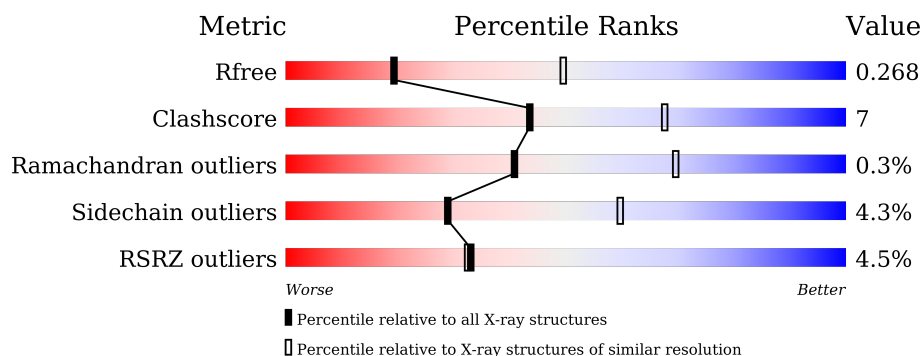
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.94 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	2969 (2.98-2.90)
Clashscore	141614	3218 (2.98-2.90)
Ramachandran outliers	138981	3122 (2.98-2.90)
Sidechain outliers	138945	3124 (2.98-2.90)
RSRZ outliers	127900	2902 (2.98-2.90)

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 $\leq 5\%$. 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	150	<div> <div style="width: 100%; height: 10px; background-color: red;"></div> <div style="display: flex; justify-content: space-between; align-items: center;"> % <div style="width: 100%; height: 10px; background-color: green;"></div> </div> <div style="display: flex; justify-content: space-between; align-items: center;"> 62% 15% • 22% </div> </div>
2	B	213	<div> <div style="width: 100%; height: 10px; background-color: red;"></div> <div style="display: flex; justify-content: space-between; align-items: center;"> 6% <div style="width: 100%; height: 10px; background-color: green;"></div> </div> <div style="display: flex; justify-content: space-between; align-items: center;"> 75% 16% • 8% </div> </div>
3	C	3	<div> <div style="width: 100%; height: 10px; background-color: green;"></div> <div style="display: flex; justify-content: space-between; align-items: center;"> 33% 67% </div> </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2509 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 sybody MR17-K99Y.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	117	Total	C	N	O	S	0	2	0
			907	565	161	177	4			

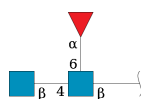
- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	195	Total	C	N	O	S	0	4	0
			1558	999	257	294	8			

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	327	ALA	-	expression tag	UNP P0DTC2
B	328	GLY	-	expression tag	UNP P0DTC2
B	329	SER	-	expression tag	UNP P0DTC2
B	532	GLY	-	expression tag	UNP P0DTC2
B	533	THR	-	expression tag	UNP P0DTC2
B	534	LEU	-	expression tag	UNP P0DTC2
B	535	GLU	-	expression tag	UNP P0DTC2
B	536	VAL	-	expression tag	UNP P0DTC2
B	537	LEU	-	expression tag	UNP P0DTC2
B	538	PHE	-	expression tag	UNP P0DTC2
B	539	GLN	-	expression tag	UNP P0DTC2

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	3	Total	C	N	O	0	0	0
			38	22	2	14			

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			6	3	3		

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	74.19Å 74.19Å 158.40Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.90 – 2.94 49.90 – 2.94	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.90-2.94) 99.9 (49.90-2.94)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.61 (at 2.96Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.214 , 0.267 0.214 , 0.268	Depositor DCC
R_{free} test set	596 reflections (5.29%)	wwPDB-VP
Wilson B-factor (Å ²)	85.4	Xtriage
Anisotropy	0.277	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 50.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.035 for -h,-k,l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	2509	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.76% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: NAG, GOL, 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	0.27	0/930	0.51	0/1256
2	B	0.27	0/1614	0.46	0/2198
All	All	0.27	0/2544	0.48	0/3454

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	907	0	874	16	0
2	B	1558	0	1473	18	0
3	C	38	0	34	0	0
4	B	6	0	8	1	0
All	All	2509	0	2389	33	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 7.

All (33) 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:91:THR:HG22	1:A:118:THR:HA	1.62	0.82
1:A:88:PRO:HA	1:A:119:VAL:HG13	1.69	0.74
1:A:91:THR:HG23	1:A:119:VAL:HG12	1.71	0.71
1:A:4:LEU:HD11	1:A:98:VAL:HG23	1.74	0.69
2:B:375:SER:HB3	2:B:436:TRP:HA	1.75	0.68
1:A:91:THR:CG2	1:A:119:VAL:H	2.09	0.66
1:A:59:ARG:NH2	2:B:488:CYS:O	2.31	0.63
2:B:412:PRO:HG3	2:B:429:PHE:HB3	1.85	0.58
1:A:91:THR:HG22	1:A:119:VAL:H	1.70	0.57
2:B:364:ASP:O	2:B:367:VAL:HG12	2.06	0.56
1:A:12:VAL:O	1:A:119:VAL:HA	2.06	0.55
2:B:458:LYS:HD2	2:B:459:SER:N	2.21	0.55
2:B:359:SER:HA	2:B:524:VAL:HG13	1.90	0.54
1:A:2:VAL:HG11	1:A:104:LEU:HD13	1.89	0.54
1:A:83:MET:HB3	1:A:86:LEU:HD21	1.90	0.54
2:B:461:LEU:HD22	2:B:465:GLU:HB3	1.92	0.51
1:A:28:PRO:O	1:A:31:VAL:HG22	2.11	0.50
2:B:472:ILE:H	2:B:472:ILE:HD12	1.79	0.48
1:A:97:ASN:OD1	1:A:106:TYR:OH	2.32	0.48
2:B:502:GLY:O	2:B:506:GLN:HG3	2.16	0.46
2:B:520:ALA:HB1	2:B:521:PRO:HD2	1.98	0.45
2:B:395:VAL:HG22	2:B:515:PHE:HD1	1.81	0.45
1:A:40:ALA:HB3	1:A:43:LYS:HD3	1.98	0.44
2:B:339:GLY:O	2:B:343:ASN:HB2	2.17	0.44
1:A:27:PHE:CZ	1:A:98:VAL:HG21	2.53	0.44
1:A:14:ALA:HA	1:A:86:LEU:O	2.20	0.42
2:B:354:ASN:HD22	2:B:354:ASN:HA	1.64	0.42
2:B:372:ALA:CB	4:B:604:GOL:H2	2.50	0.41
2:B:438:SER:HB3	2:B:509:ARG:HG3	2.03	0.41
2:B:346:ARG:NH1	2:B:354:ASN:OD1	2.54	0.41
2:B:473:TYR:CE2	2:B:475:ALA:HB2	2.56	0.41
1:A:18:LEU:HD23	1:A:18:LEU:HA	1.87	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	115/150 (77%)	113 (98%)	2 (2%)	0	100	100
2	B	197/213 (92%)	186 (94%)	10 (5%)	1 (0%)	29	60
All	All	312/363 (86%)	299 (96%)	12 (4%)	1 (0%)	41	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	481	ASN

5.3.2 Protein sidechains ⓘ

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	94/120 (78%)	90 (96%)	4 (4%)	29	60
2	B	170/183 (93%)	162 (95%)	8 (5%)	26	57
All	All	264/303 (87%)	252 (96%)	12 (4%)	29	58

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	17	SER
1	A	43	LYS
1	A	45	ARG
1	A	120	SER
2	B	347	PHE
2	B	373	SER
2	B	383	SER
2	B	424	LYS
2	B	458	LYS
2	B	484[A]	GLU
2	B	484[B]	GLU
2	B	493	GLN

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

Mol	Chain	Res	Type
1	A	1	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 ⓘ

3 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	$\# Z > 2$
3	NAG	C	1	2,3	14,14,15	0.33	0	17,19,21	0.69	1 (5%)
3	NAG	C	2	3	14,14,15	0.30	0	17,19,21	0.37	0
3	FUC	C	3	3	10,10,11	1.70	2 (20%)	14,14,16	1.80	4 (28%)

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
3	NAG	C	1	2,3	-	4/6/23/26	0/1/1/1
3	NAG	C	2	3	-	1/6/23/26	0/1/1/1
3	FUC	C	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	C	3	FUC	C1-C2	3.80	1.60	1.52
3	C	3	FUC	O5-C1	3.05	1.48	1.43

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	3	FUC	C1-O5-C5	3.98	121.80	112.78
3	C	3	FUC	C1-C2-C3	3.39	113.83	109.67
3	C	3	FUC	O5-C1-C2	2.56	114.72	110.77
3	C	1	NAG	C1-O5-C5	2.23	115.21	112.19
3	C	3	FUC	O5-C5-C4	2.13	113.34	109.52

There are no chirality outliers.

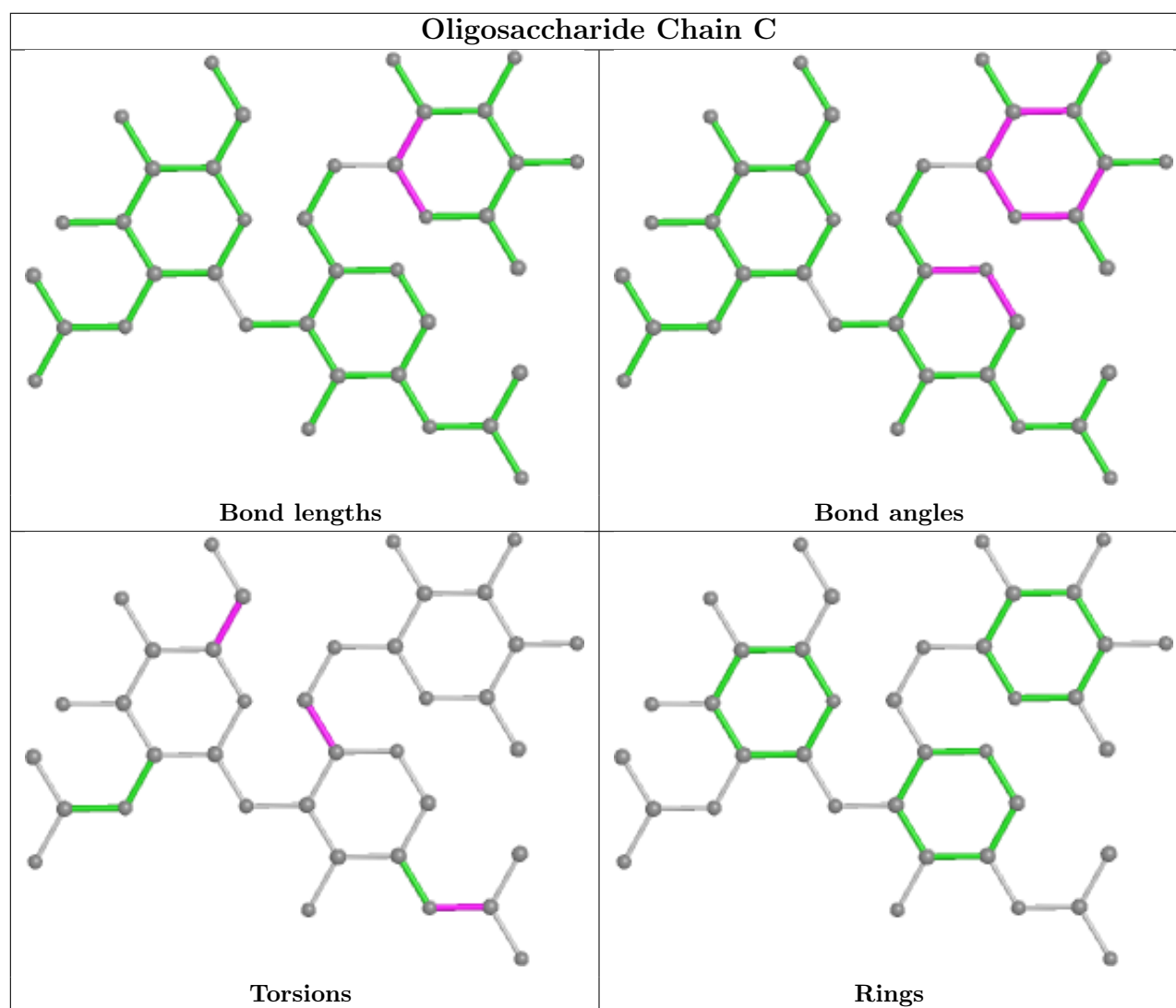
All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	1	NAG	O5-C5-C6-O6
3	C	1	NAG	C4-C5-C6-O6
3	C	1	NAG	C8-C7-N2-C2
3	C	1	NAG	O7-C7-N2-C2
3	C	2	NAG	O5-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.



5.6 Ligand geometry [i](#)

1 ligand is 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	$\# Z > 2$
4	GOL	B	604	-	5,5,5	0.87	0	5,5,5	0.99	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
4	GOL	B	604	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	604	GOL	O1-C1-C2-C3
4	B	604	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	604	GOL	1	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 [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	117/150 (78%)	0.13	1 (0%) 84 85	55, 73, 108, 129	0
2	B	195/213 (91%)	0.26	13 (6%) 17 16	57, 75, 114, 145	0
All	All	312/363 (85%)	0.21	14 (4%) 33 32	55, 74, 113, 145	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	519	HIS	4.1
2	B	518	LEU	3.8
2	B	334	ASN	3.1
2	B	520	ALA	3.1
2	B	333	THR	3.0
2	B	481	ASN	2.6
2	B	521	PRO	2.6
1	A	95	TYR	2.4
2	B	517	LEU	2.3
2	B	335	LEU	2.3
2	B	477	SER	2.2
2	B	390	LEU	2.1
2	B	525	CYS	2.0
2	B	523	THR	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.

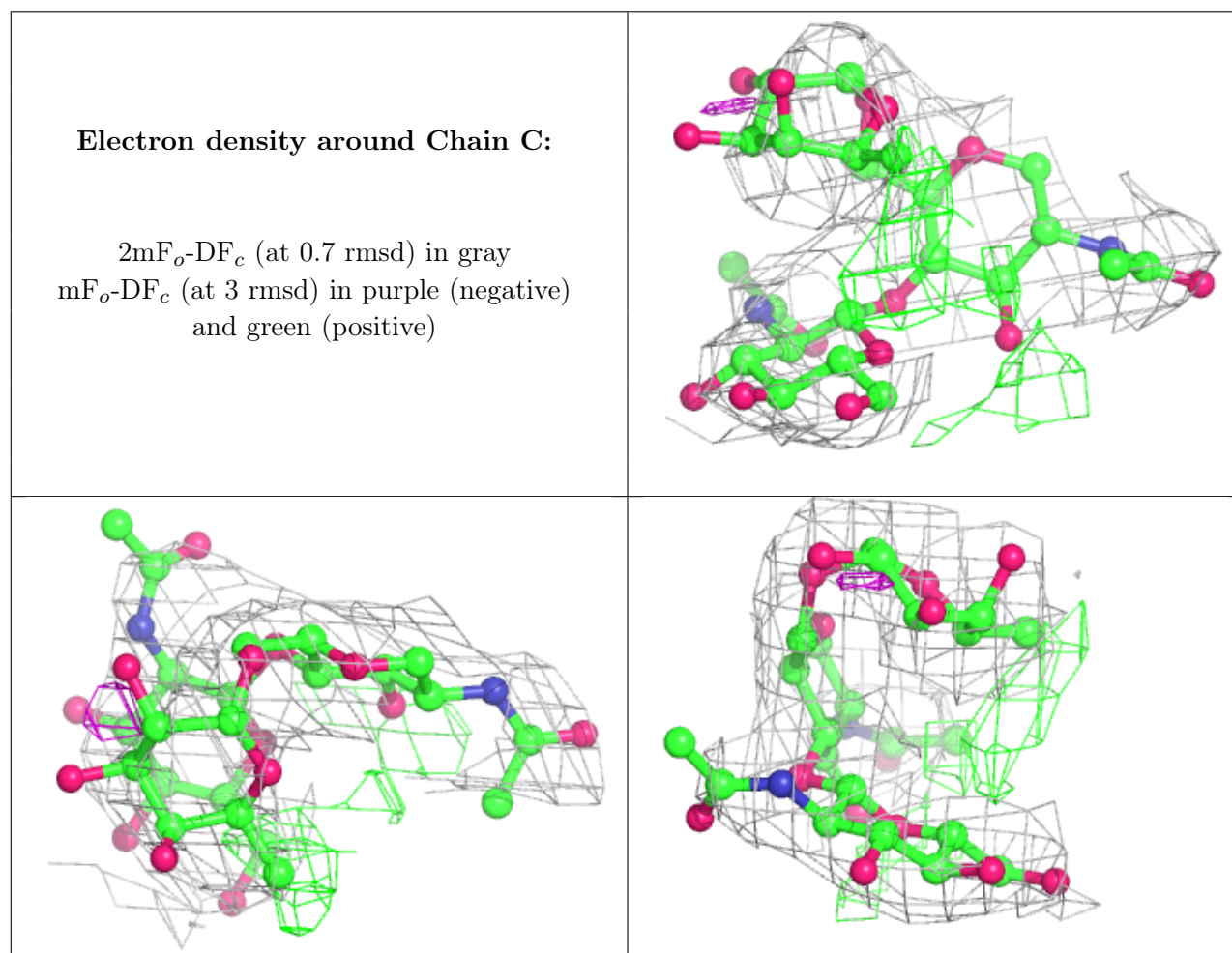
6.3 Carbohydrates [i](#)

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, 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	FUC	C	3	10/11	0.73	0.36	135,151,157,159	0
3	NAG	C	1	14/15	0.86	0.17	106,118,137,141	0
3	NAG	C	2	14/15	0.89	0.20	118,133,137,138	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 ⓘ

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(\AA^2)	Q<0.9
4	GOL	B	604	6/6	0.90	0.25	73,86,87,93	0

6.5 Other polymers [i](#)

There are no such residues in this entry.