



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 17, 2023 – 03:16 PM EDT

PDB ID : 4ZMJ  
Title : Crystal Structure of Ligand-Free BG505 SOSIP.664 HIV-1 Env Trimer  
Authors : Kwon, Y.D.; Kwong, P.D.  
Deposited on : 2015-05-04  
Resolution : 3.31 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](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 : 2.35.1  
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) : 2.35.1

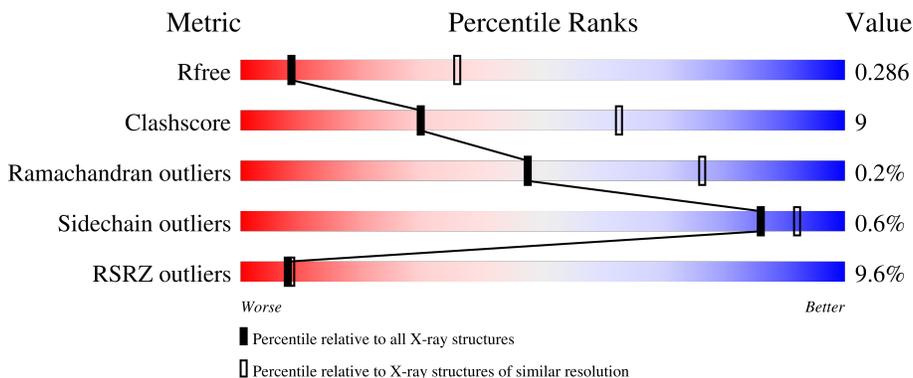
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.31 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	1089 (3.36-3.28)
Clashscore	141614	1137 (3.36-3.28)
Ramachandran outliers	138981	1115 (3.36-3.28)
Sidechain outliers	138945	1114 (3.36-3.28)
RSRZ outliers	127900	1059 (3.36-3.28)

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  $\geq 3$ , 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	G	481	 10% 68% 26% 6%
2	B	153	 6% 64% 16% 20%
3	A	6	 17% 83%
4	C	2	 50% 50%
4	D	2	 100%

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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
5	NAG	B	701	-	-	-	X
5	NAG	B	702	-	-	-	X
5	NAG	G	601	-	-	-	X
5	NAG	G	610	-	-	-	X
5	NAG	G	620	-	-	-	X

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 4856 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 glycoprotein gp160.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	G	450	3543	2224	626	665	28	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	332	ASN	THR	engineered mutation	UNP Q2N0S6
G	501	CYS	ALA	engineered mutation	UNP Q2N0S6
G	508	ARG	-	expression tag	UNP Q2N0S6
G	509	ARG	-	expression tag	UNP Q2N0S6
G	510	ARG	-	expression tag	UNP Q2N0S6
G	511	ARG	-	expression tag	UNP Q2N0S6
G	512	ARG	-	expression tag	UNP Q2N0S6
G	513	ARG	-	expression tag	UNP Q2N0S6

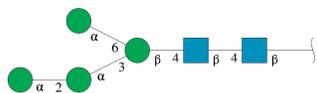
- Molecule 2 is a protein called Envelope glycoprotein gp160.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	123	975	613	169	187	6	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	559	PRO	ILE	engineered mutation	UNP Q2N0S6
B	605	CYS	THR	engineered mutation	UNP Q2N0S6

- Molecule 3 is an oligosaccharide called alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[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.



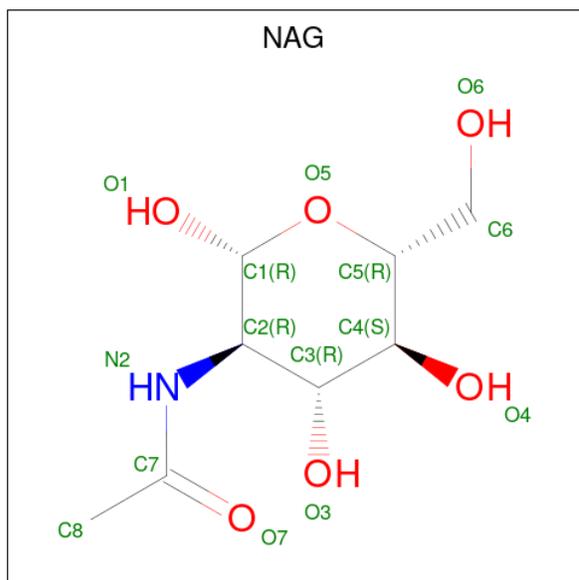
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	A	6	72	40	2	30	0	0	0

- Molecule 4 is an oligosaccharide called 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			
4	C	2	28	16	2	10	0	0	0
4	D	2	28	16	2	10	0	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	G	1	14	8	1	5	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		





- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucofuranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:  50% 50%



- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucofuranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:  100%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.18Å 107.18Å 103.06Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	35.08 – 3.31 47.55 – 3.31	Depositor EDS
% Data completeness (in resolution range)	63.7 (35.08-3.31) 63.8 (47.55-3.31)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.49 (at 3.33Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.266 , 0.285 0.269 , 0.286	Depositor DCC
$R_{free}$ test set	323 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	88.7	Xtrriage
Anisotropy	0.026	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 42.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	0.099 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	4856	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	95.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.15% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, MAN, BMA

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	G	0.29	0/3617	0.61	0/4911
2	B	0.22	0/992	0.38	0/1345
All	All	0.28	0/4609	0.57	0/6256

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	G	0	4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	354	GLY	Peptide
1	G	431	GLY	Peptide
1	G	494	LEU	Peptide
1	G	57	ASP	Peptide

### 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	G	3543	0	3481	79	2
2	B	975	0	946	18	1
3	A	72	0	61	3	0
4	C	28	0	25	1	0
4	D	28	0	25	0	0
5	B	42	0	39	0	0
5	G	168	0	156	1	1
All	All	4856	0	4733	90	2

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 9.

All (90) 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:G:270:VAL:HG23	1:G:348:GLN:HG3	1.64	0.78
1:G:301:ASN:HB3	1:G:323:ILE:HB	1.64	0.77
1:G:270:VAL:HG12	1:G:288:PHE:HA	1.69	0.74
1:G:50:THR:O	1:G:103:GLN:NE2	2.16	0.73
1:G:138:ILE:HG23	1:G:139:THR:HG23	1.71	0.72
1:G:386:ASN:HB3	1:G:417:PRO:HD2	1.75	0.69
1:G:37:THR:HG22	2:B:605:CYS:HA	1.74	0.68
1:G:74:CYS:SG	2:B:571:TRP:NE1	2.65	0.68
2:B:598:CYS:HB3	2:B:604:CYS:HB3	1.74	0.68
1:G:215:ILE:HD12	1:G:253:PRO:HD3	1.79	0.63
1:G:425:ASN:ND2	1:G:429:ARG:O	2.31	0.63
1:G:221:ALA:HB3	2:B:582:ALA:HB1	1.79	0.63
1:G:36:VAL:O	2:B:606:THR:OG1	2.15	0.62
3:A:1:NAG:H61	3:A:2:NAG:HN2	1.63	0.62
1:G:295:ASN:O	1:G:332:ASN:N	2.27	0.62
1:G:195:ASN:HD22	1:G:423:ILE:HG21	1.65	0.61
1:G:50:THR:HG22	1:G:488:VAL:HG11	1.82	0.61
1:G:55:ALA:HB1	1:G:77:THR:HA	1.83	0.60
2:B:593:LEU:HD21	2:B:601:LYS:HA	1.83	0.60
1:G:57:ASP:N	1:G:77:THR:OG1	2.36	0.59
1:G:330:HIS:HB2	1:G:415:THR:HG23	1.87	0.57
1:G:113:ASP:OD1	1:G:429:ARG:NH2	2.37	0.57
1:G:499:THR:HG23	1:G:501:CYS:H	1.68	0.57
1:G:385:CYS:HA	1:G:418:CYS:HA	1.87	0.56
1:G:421:LYS:HE3	1:G:424:ILE:HG22	1.88	0.56
1:G:38:VAL:HG22	1:G:496:VAL:HG12	1.86	0.55
5:G:608:NAG:H83	5:G:608:NAG:H3	1.87	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:152:GLY:HA3	1:G:178:ARG:HH21	1.71	0.55
1:G:272:ILE:HD13	1:G:349:LEU:HD23	1.88	0.55
1:G:101:VAL:HG13	1:G:479:TRP:HB2	1.89	0.54
2:B:649:SER:O	2:B:653:GLN:NE2	2.42	0.54
1:G:45:TRP:H	2:B:629:LEU:HD21	1.72	0.53
2:B:650:GLN:O	2:B:654:GLU:N	2.33	0.53
1:G:75:VAL:HG13	1:G:76:PRO:HD2	1.89	0.53
1:G:270:VAL:HG13	1:G:344:LYS:HE3	1.91	0.52
1:G:296:CYS:HA	1:G:331:CYS:HA	1.91	0.51
1:G:298:ARG:HB3	1:G:443:ILE:HB	1.93	0.50
1:G:384:TYR:O	1:G:419:ARG:N	2.41	0.50
2:B:570:VAL:HG12	2:B:571:TRP:H	1.76	0.49
1:G:161:MET:O	1:G:170:GLN:N	2.42	0.49
1:G:498:PRO:HB3	2:B:610:TRP:CD2	2.48	0.49
1:G:175:LEU:HB2	1:G:320:THR:HB	1.94	0.49
1:G:231:LYS:HG2	1:G:267:GLU:OE1	2.13	0.49
1:G:83:GLU:HA	1:G:245:VAL:HG12	1.96	0.48
1:G:390:LEU:HG	1:G:416:LEU:HD21	1.95	0.48
1:G:95:MET:SD	1:G:273:ARG:HD3	2.53	0.48
1:G:220:PRO:HG3	2:B:578:ALA:HB1	1.95	0.48
1:G:138:ILE:HA	1:G:139:THR:HA	1.55	0.48
1:G:122:LEU:HB2	1:G:201:ILE:HG22	1.95	0.47
2:B:636:SER:O	2:B:639:THR:OG1	2.29	0.47
1:G:153:GLU:HA	1:G:419:ARG:NH2	2.29	0.46
1:G:350:ARG:CZ	1:G:397:SER:HA	2.45	0.46
2:B:646:LEU:O	2:B:650:GLN:HB2	2.16	0.46
3:A:1:NAG:H61	3:A:2:NAG:N2	2.29	0.46
1:G:212:PRO:HG2	1:G:254:VAL:HG22	1.96	0.46
1:G:95:MET:SD	1:G:235:GLY:HA3	2.56	0.46
1:G:227:LYS:HA	1:G:485:LYS:O	2.16	0.45
1:G:477:ASP:OD1	1:G:480:ARG:NH1	2.48	0.45
1:G:53:PHE:HB2	2:B:575:GLN:HE21	1.82	0.45
1:G:153:GLU:HA	1:G:419:ARG:HH22	1.82	0.45
1:G:117:LYS:HB3	1:G:118:PRO:HD3	2.00	0.44
1:G:178:ARG:HB3	1:G:419:ARG:HH21	1.81	0.44
1:G:102:GLU:O	1:G:106:THR:HG23	2.18	0.44
1:G:227:LYS:HB3	1:G:243:SER:OG	2.18	0.44
1:G:54:CYS:SG	2:B:571:TRP:CZ2	3.10	0.44
1:G:134:VAL:HG12	1:G:136:ASN:H	1.82	0.44
1:G:212:PRO:HB2	1:G:252:LYS:HB3	1.99	0.44
1:G:122:LEU:HD13	1:G:125:LEU:HD12	2.00	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:276:ASN:HB3	1:G:279:ASN:HB3	1.99	0.43
1:G:164:GLU:HB2	1:G:165:LEU:HD12	2.01	0.43
1:G:131:CYS:HA	1:G:156:ASN:O	2.18	0.43
1:G:426:MET:HB2	1:G:427:TRP:CE3	2.53	0.43
2:B:573:ILE:HG13	2:B:576:LEU:HD23	2.00	0.43
1:G:162:THR:HG22	1:G:163:THR:O	2.17	0.43
1:G:334:SER:HB3	1:G:337:THR:HB	2.00	0.43
1:G:56:SER:O	1:G:56:SER:OG	2.37	0.43
1:G:493:PRO:HG3	2:B:544:LEU:HD21	2.00	0.42
1:G:270:VAL:HG22	1:G:344:LYS:HG2	2.00	0.42
1:G:303:THR:HB	1:G:321:GLY:HA3	2.02	0.42
1:G:211:GLU:OE2	3:A:2:NAG:H5	2.19	0.42
1:G:212:PRO:HB3	1:G:253:PRO:HD2	2.02	0.41
1:G:65:LYS:HB2	1:G:68:VAL:HG13	2.02	0.41
1:G:198:THR:OG1	1:G:199:SER:N	2.53	0.41
1:G:332:ASN:HA	1:G:414:ILE:O	2.19	0.41
1:G:492:GLU:HA	1:G:493:PRO:HD2	1.91	0.41
1:G:175:LEU:HD11	1:G:321(A):ASP:HA	2.02	0.41
1:G:42:VAL:HG22	1:G:493:PRO:O	2.21	0.41
1:G:197:ASN:HD22	4:C:1:NAG:H83	1.85	0.40
1:G:39:TYR:HE2	1:G:497:ALA:HB3	1.86	0.40
1:G:216:HIS:HB3	1:G:247:CYS:HB3	2.02	0.40

All (2) 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:G:136:ASN:ND2	5:G:610:NAG:O3[5_564]	2.16	0.04
1:G:187:SER:OG	2:B:660:LEU:O[2_774]	2.16	0.04

## 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	G	444/481 (92%)	411 (93%)	32 (7%)	1 (0%)	47	76
2	B	119/153 (78%)	105 (88%)	14 (12%)	0	100	100
All	All	563/634 (89%)	516 (92%)	46 (8%)	1 (0%)	47	76

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	57	ASP

### 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	G	402/428 (94%)	401 (100%)	1 (0%)	93	97
2	B	105/129 (81%)	103 (98%)	2 (2%)	57	78
All	All	507/557 (91%)	504 (99%)	3 (1%)	86	92

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	138	ILE
2	B	598	CYS
2	B	625	ASN

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

Mol	Chain	Res	Type
2	B	575	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](#)

10 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	A	1	3,1	14,14,15	0.23	0	17,19,21	0.58	0
3	NAG	A	2	3	14,14,15	0.31	0	17,19,21	0.39	0
3	BMA	A	3	3	11,11,12	0.68	0	15,15,17	0.78	0
3	MAN	A	4	3	11,11,12	0.73	0	15,15,17	1.00	1 (6%)
3	MAN	A	5	3	11,11,12	0.75	0	15,15,17	1.27	2 (13%)
3	MAN	A	6	3	11,11,12	0.62	0	15,15,17	1.16	2 (13%)
4	NAG	C	1	4,1	14,14,15	0.19	0	17,19,21	0.46	0
4	NAG	C	2	4	14,14,15	0.23	0	17,19,21	0.40	0
4	NAG	D	1	4,1	14,14,15	0.29	0	17,19,21	0.41	0
4	NAG	D	2	4	14,14,15	0.31	0	17,19,21	0.55	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
3	NAG	A	1	3,1	-	2/6/23/26	0/1/1/1
3	NAG	A	2	3	-	0/6/23/26	0/1/1/1
3	BMA	A	3	3	-	0/2/19/22	0/1/1/1
3	MAN	A	4	3	-	2/2/19/22	0/1/1/1
3	MAN	A	5	3	-	2/2/19/22	1/1/1/1
3	MAN	A	6	3	-	2/2/19/22	0/1/1/1
4	NAG	C	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	C	2	4	-	2/6/23/26	0/1/1/1

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	D	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	D	2	4	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	5	MAN	C1-O5-C5	3.51	116.95	112.19
3	A	6	MAN	C1-O5-C5	2.70	115.85	112.19
3	A	4	MAN	O2-C2-C3	-2.55	105.02	110.14
3	A	5	MAN	O2-C2-C3	-2.24	105.64	110.14
3	A	6	MAN	O2-C2-C3	-2.17	105.79	110.14

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	2	NAG	O5-C5-C6-O6
3	A	1	NAG	O5-C5-C6-O6
4	C	2	NAG	O5-C5-C6-O6
4	D	2	NAG	C4-C5-C6-O6
3	A	5	MAN	C4-C5-C6-O6
3	A	1	NAG	C4-C5-C6-O6
4	C	1	NAG	C8-C7-N2-C2
4	C	1	NAG	O7-C7-N2-C2
3	A	4	MAN	O5-C5-C6-O6
3	A	4	MAN	C4-C5-C6-O6
4	C	2	NAG	C4-C5-C6-O6
3	A	5	MAN	O5-C5-C6-O6
4	D	1	NAG	C4-C5-C6-O6
4	D	1	NAG	O5-C5-C6-O6
3	A	6	MAN	C4-C5-C6-O6
4	D	2	NAG	C3-C2-N2-C7
3	A	6	MAN	O5-C5-C6-O6

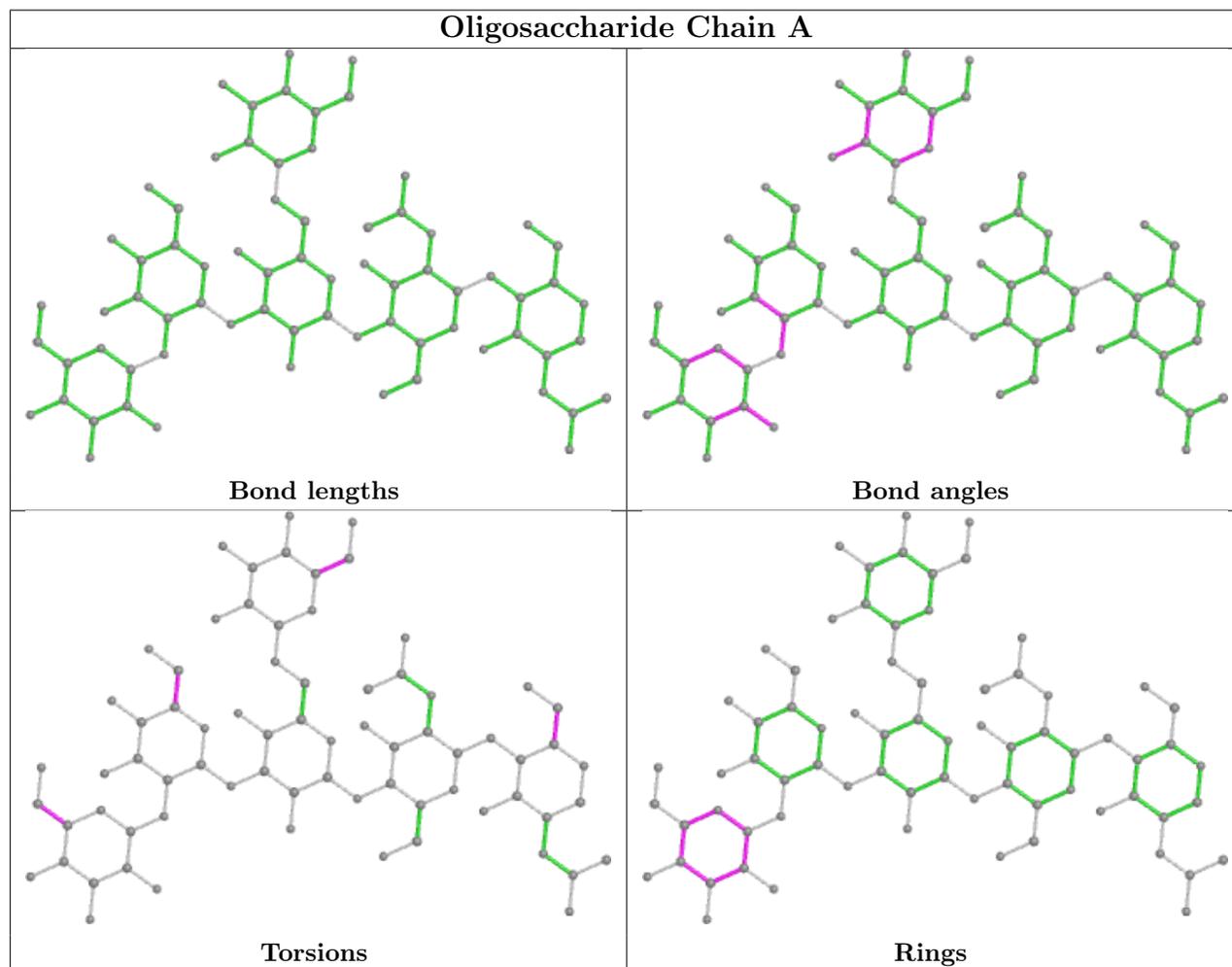
All (1) ring outliers are listed below:

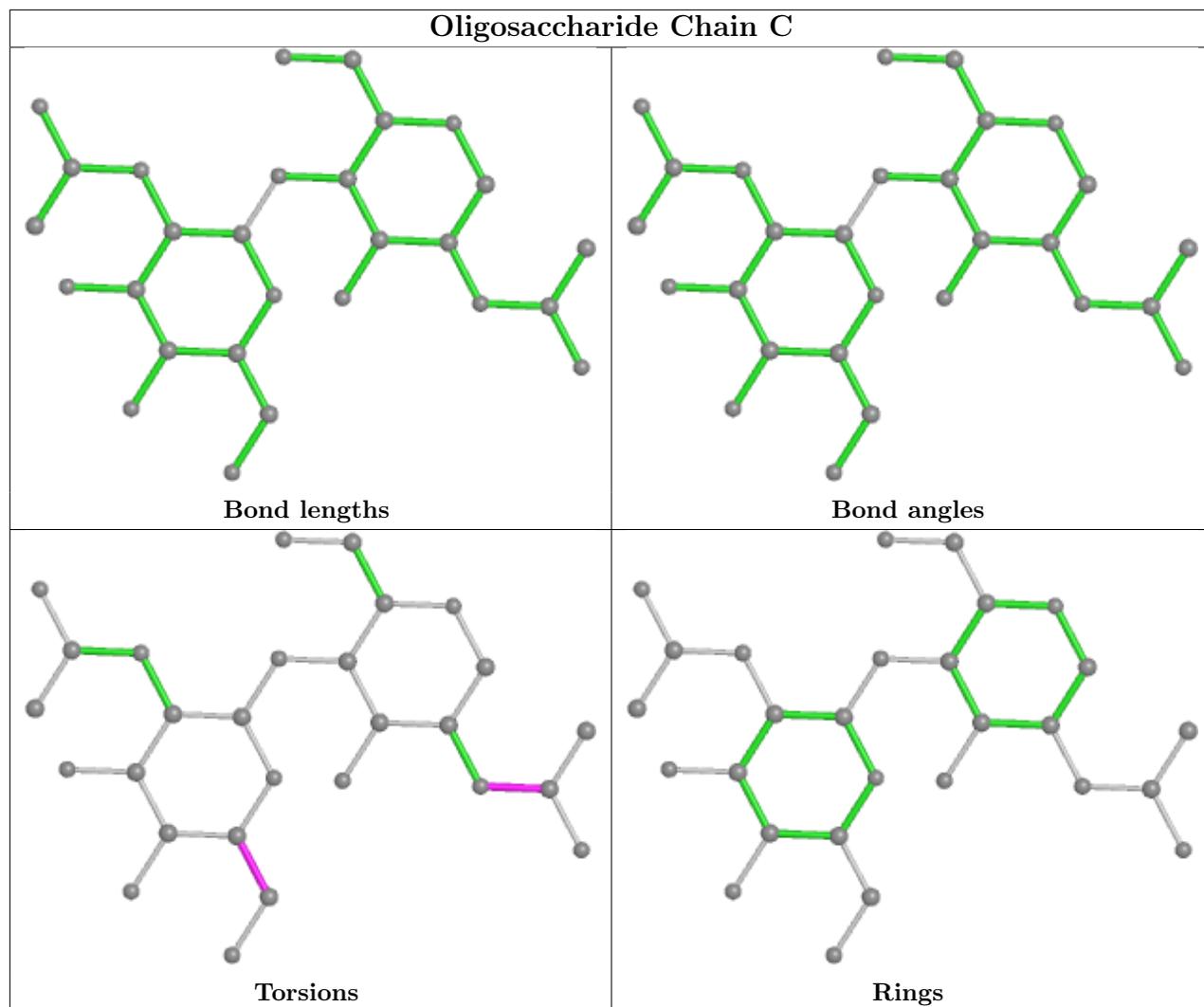
Mol	Chain	Res	Type	Atoms
3	A	5	MAN	C1-C2-C3-C4-C5-O5

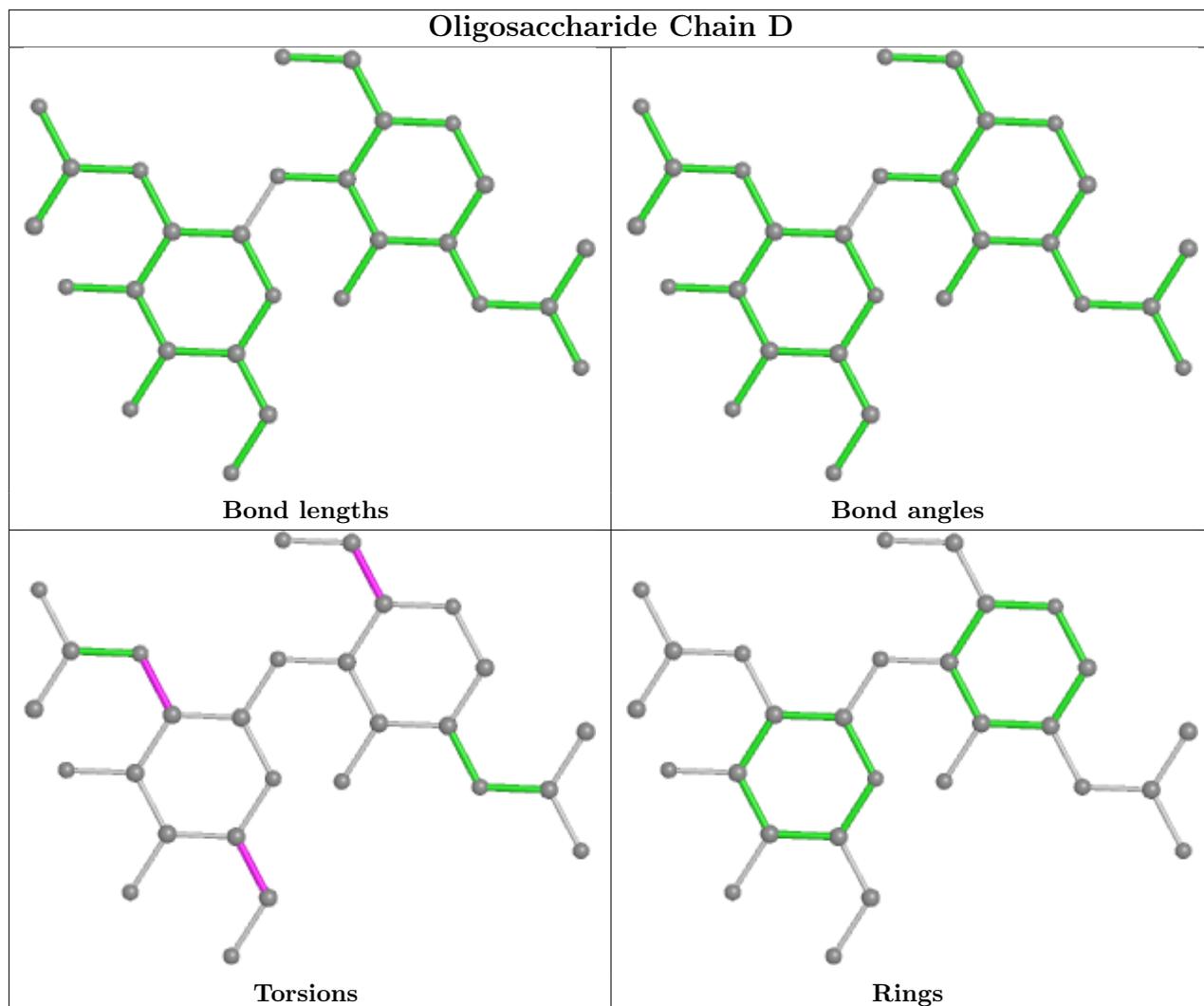
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	1	NAG	1	0
3	A	1	NAG	2	0
3	A	2	NAG	3	0

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](#)

15 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	$\# Z  > 2$
5	NAG	G	601	1	14,14,15	0.19	0	17,19,21	0.45	0
5	NAG	G	619	1	14,14,15	0.22	0	17,19,21	0.40	0
5	NAG	G	620	1	14,14,15	0.30	0	17,19,21	0.80	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	G	614	1	14,14,15	0.23	0	17,19,21	0.47	0
5	NAG	G	616	1	14,14,15	0.25	0	17,19,21	0.41	0
5	NAG	G	608	1	14,14,15	0.46	0	17,19,21	1.24	1 (5%)
5	NAG	G	615	1	14,14,15	0.21	0	17,19,21	0.40	0
5	NAG	G	610	1	14,14,15	0.38	0	17,19,21	0.57	0
5	NAG	G	613	1	14,14,15	0.34	0	17,19,21	0.46	0
5	NAG	B	701	2	14,14,15	0.19	0	17,19,21	0.45	0
5	NAG	G	609	1	14,14,15	0.27	0	17,19,21	0.37	0
5	NAG	B	703	2	14,14,15	0.29	0	17,19,21	0.36	0
5	NAG	B	702	2	14,14,15	0.23	0	17,19,21	0.34	0
5	NAG	G	611	1	14,14,15	0.20	0	17,19,21	0.41	0
5	NAG	G	612	1	14,14,15	0.23	0	17,19,21	0.57	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
5	NAG	G	601	1	-	0/6/23/26	0/1/1/1
5	NAG	G	619	1	-	2/6/23/26	0/1/1/1
5	NAG	G	620	1	-	2/6/23/26	0/1/1/1
5	NAG	G	614	1	-	2/6/23/26	0/1/1/1
5	NAG	G	616	1	-	3/6/23/26	0/1/1/1
5	NAG	G	608	1	-	5/6/23/26	0/1/1/1
5	NAG	G	615	1	-	0/6/23/26	0/1/1/1
5	NAG	G	610	1	-	3/6/23/26	0/1/1/1
5	NAG	G	613	1	-	2/6/23/26	0/1/1/1
5	NAG	B	701	2	-	2/6/23/26	0/1/1/1
5	NAG	G	609	1	-	0/6/23/26	0/1/1/1
5	NAG	B	703	2	-	0/6/23/26	0/1/1/1
5	NAG	B	702	2	-	2/6/23/26	0/1/1/1
5	NAG	G	611	1	-	2/6/23/26	0/1/1/1
5	NAG	G	612	1	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
5	G	608	NAG	C2-N2-C7	4.29	129.01	122.90
5	G	620	NAG	C1-O5-C5	2.74	115.90	112.19

There are no chirality outliers.

All (28) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	G	620	NAG	O5-C5-C6-O6
5	G	611	NAG	O5-C5-C6-O6
5	G	611	NAG	C4-C5-C6-O6
5	B	701	NAG	O5-C5-C6-O6
5	B	702	NAG	O5-C5-C6-O6
5	G	619	NAG	O5-C5-C6-O6
5	G	612	NAG	O5-C5-C6-O6
5	G	620	NAG	C4-C5-C6-O6
5	G	608	NAG	C8-C7-N2-C2
5	G	608	NAG	O7-C7-N2-C2
5	G	610	NAG	C8-C7-N2-C2
5	G	610	NAG	O7-C7-N2-C2
5	G	614	NAG	C8-C7-N2-C2
5	G	614	NAG	O7-C7-N2-C2
5	G	616	NAG	C8-C7-N2-C2
5	G	616	NAG	O7-C7-N2-C2
5	G	612	NAG	C4-C5-C6-O6
5	G	619	NAG	C4-C5-C6-O6
5	B	702	NAG	C4-C5-C6-O6
5	G	613	NAG	O5-C5-C6-O6
5	G	608	NAG	O5-C5-C6-O6
5	B	701	NAG	C4-C5-C6-O6
5	G	613	NAG	C4-C5-C6-O6
5	G	608	NAG	C4-C5-C6-O6
5	G	612	NAG	C3-C2-N2-C7
5	G	610	NAG	C4-C5-C6-O6
5	G	608	NAG	C3-C2-N2-C7
5	G	616	NAG	C4-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	608	NAG	1	0
5	G	610	NAG	0	1

## 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<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	G	450/481 (93%)	0.68	46 (10%) <b>6</b> <b>7</b>	40, 84, 145, 192	0
2	B	123/153 (80%)	0.61	9 (7%) <b>15</b> <b>15</b>	42, 103, 144, 179	0
All	All	573/634 (90%)	0.66	55 (9%) <b>8</b> <b>8</b>	40, 89, 146, 192	0

All (55) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	619	LEU	6.2
1	G	411	ASN	5.9
1	G	231	LYS	5.4
1	G	136	ASN	5.3
1	G	72	HIS	5.1
1	G	66	HIS	5.0
1	G	413	SER	4.4
1	G	232	LYS	4.4
2	B	618	ASN	4.3
2	B	625	ASN	4.1
1	G	505	VAL	4.1
1	G	267	GLU	3.9
2	B	547	GLY	3.9
1	G	395	TRP	3.7
1	G	396	ILE	3.6
1	G	352	HIS	3.5
1	G	134	VAL	3.4
1	G	463	SER	3.4
1	G	353	PHE	3.4
1	G	67	ASN	3.2
1	G	230	ASP	3.2
1	G	137	ASN	3.0
1	G	130	GLN	3.0
1	G	265	LEU	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	G	188	ASN	2.9
1	G	333	VAL	2.9
1	G	412	ASP	2.8
1	G	259	LEU	2.8
1	G	455	THR	2.8
1	G	135	THR	2.7
1	G	187	SER	2.7
1	G	65	LYS	2.6
2	B	543	ASN	2.5
1	G	138	ILE	2.4
1	G	73	ALA	2.4
2	B	614	TRP	2.3
1	G	140	ASP	2.3
1	G	426	MET	2.3
2	B	571	TRP	2.3
1	G	71	THR	2.3
1	G	98	ASN	2.3
1	G	153	GLU	2.2
2	B	664	ASP	2.2
1	G	359	ILE	2.2
1	G	275	GLU	2.2
1	G	427	TRP	2.2
1	G	95	MET	2.2
1	G	325	ASP	2.1
1	G	336	ALA	2.1
1	G	462	ASN	2.1
1	G	271	MET	2.1
1	G	442	VAL	2.1
2	B	602	LEU	2.1
1	G	168	LYS	2.1
1	G	90	THR	2.0

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

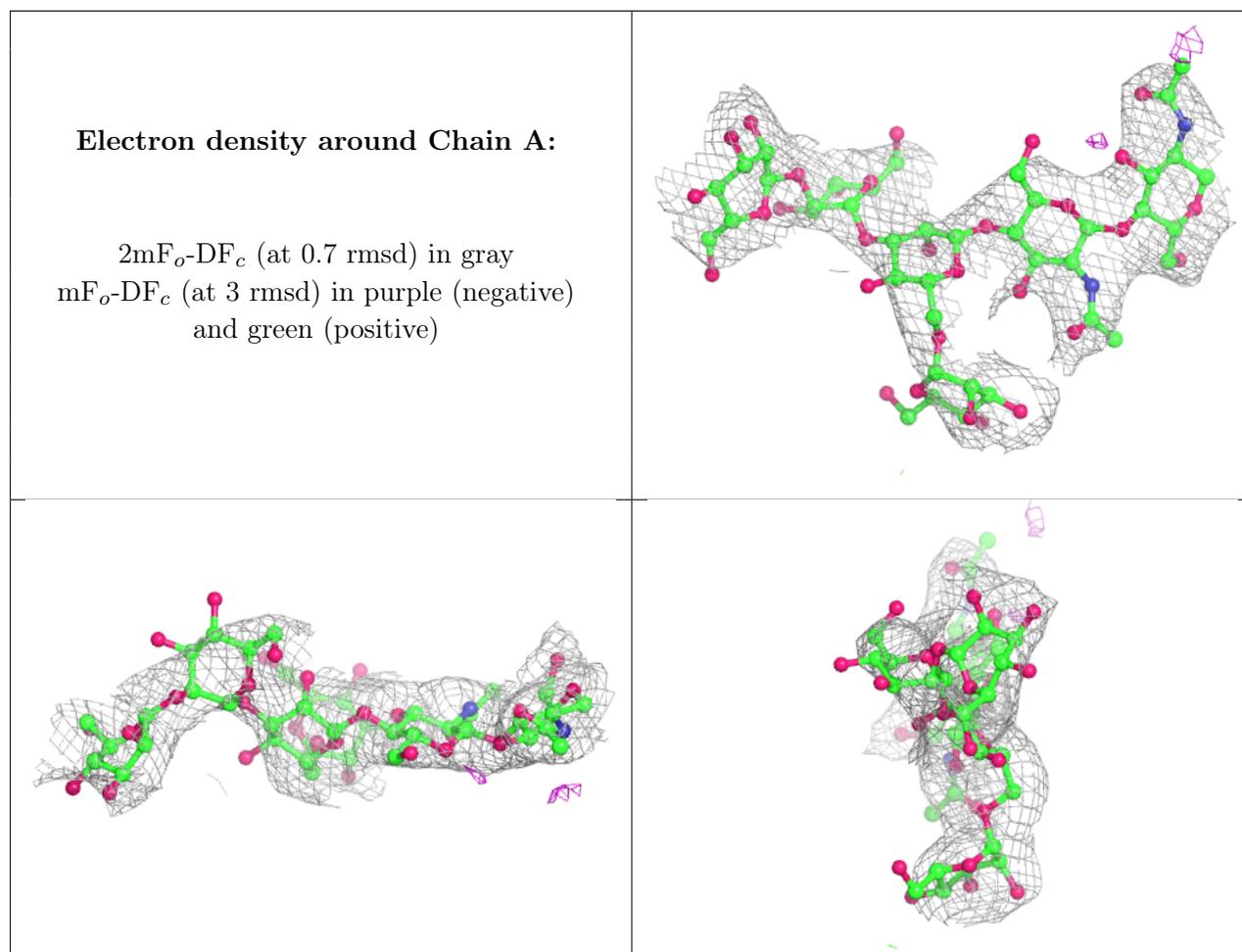
There are no non-standard protein/DNA/RNA residues in this entry.

## 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, 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.

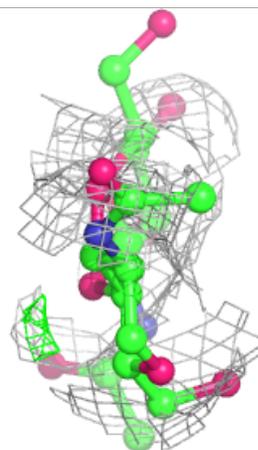
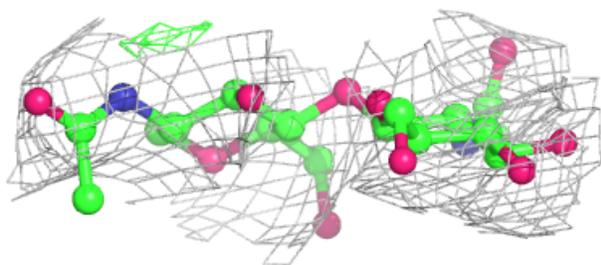
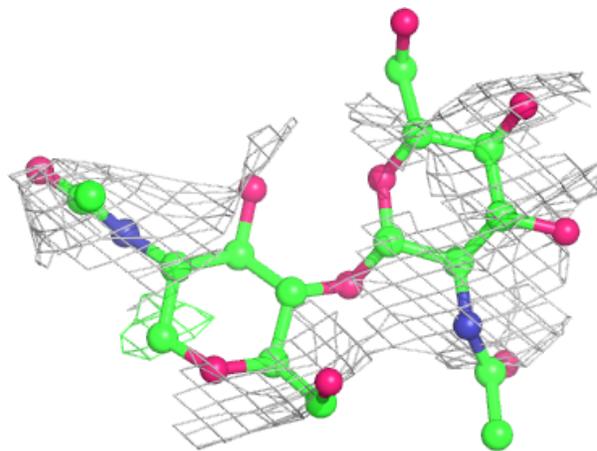
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MAN	A	6	11/12	0.69	0.33	86,135,151,161	0
3	MAN	A	5	11/12	0.71	0.28	97,119,129,134	0
4	NAG	C	2	14/15	0.72	0.35	116,141,169,181	0
4	NAG	C	1	14/15	0.73	0.27	104,126,135,137	0
3	MAN	A	4	11/12	0.73	0.34	127,157,166,167	0
3	BMA	A	3	11/12	0.76	0.23	117,131,147,151	0
4	NAG	D	1	14/15	0.82	0.30	92,103,113,120	0
4	NAG	D	2	14/15	0.84	0.52	122,141,152,159	0
3	NAG	A	1	14/15	0.91	0.19	36,60,81,81	0
3	NAG	A	2	14/15	0.91	0.19	92,102,117,125	0

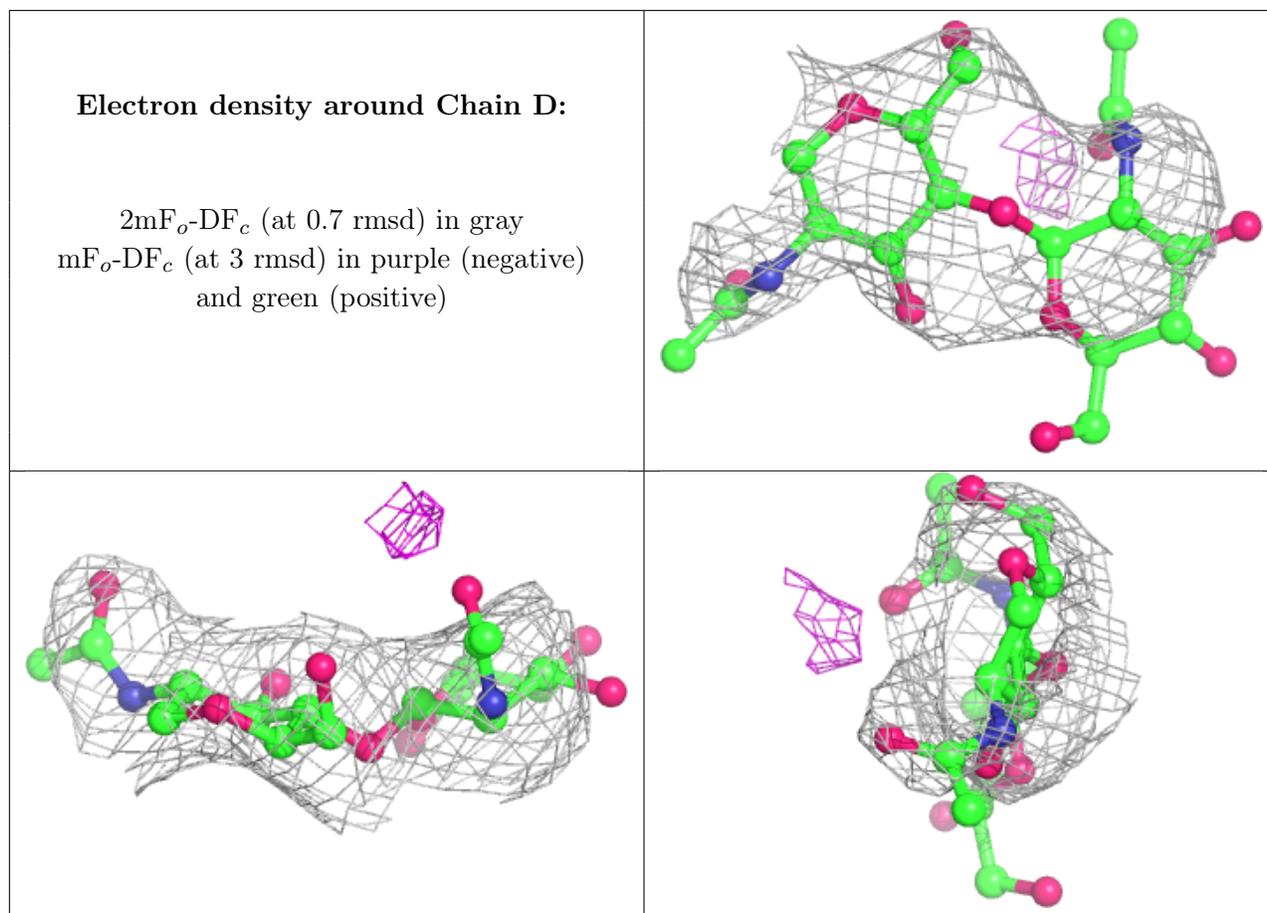
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**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)





## 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
5	NAG	B	702	14/15	0.49	0.42	121,152,162,169	0
5	NAG	G	620	14/15	0.59	0.58	112,127,141,154	0
5	NAG	G	601	14/15	0.63	0.58	85,105,115,117	0
5	NAG	B	701	14/15	0.64	0.47	144,157,164,167	0
5	NAG	G	610	14/15	0.68	0.63	139,150,159,160	0
5	NAG	B	703	14/15	0.68	0.39	121,131,147,153	0
5	NAG	G	613	14/15	0.70	0.29	93,142,161,172	0
5	NAG	G	614	14/15	0.74	0.30	79,96,111,111	0
5	NAG	G	616	14/15	0.75	0.38	102,126,135,147	0
5	NAG	G	612	14/15	0.78	0.31	109,134,167,177	0
5	NAG	G	611	14/15	0.79	0.24	82,110,127,127	0
5	NAG	G	609	14/15	0.81	0.21	64,118,130,145	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	NAG	G	615	14/15	0.82	0.44	111,122,129,131	0
5	NAG	G	608	14/15	0.85	0.38	101,119,127,138	0
5	NAG	G	619	14/15	0.90	0.23	83,102,113,117	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.