



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 13, 2026 – 07:11 PM UTC

PDB ID : 7EE5 / pdb\_00007ee5  
Title : Crystal structure of Neu5Gc bound PltC  
Authors : Liu, X.Y.; Chen, Z.; Gao, X.  
Deposited on : 2021-03-17  
Resolution : 1.24 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 : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Buster-report : wwPDB partial adaption of 1.1.7 (2018)  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

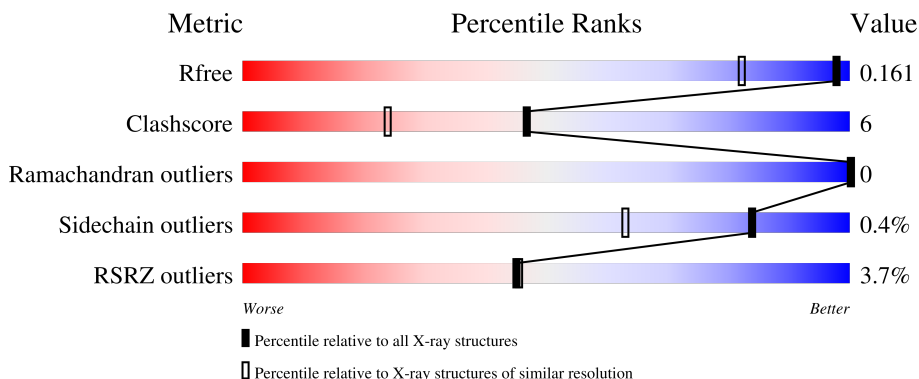
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.24 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	1630 (1.26-1.22)
Clashscore	190562	1668 (1.26-1.22)
Ramachandran outliers	187476	1635 (1.26-1.22)
Sidechain outliers	187428	1633 (1.26-1.22)
RSRZ outliers	180081	1630 (1.26-1.22)

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	A	124	 3% 83% 12% 5%
1	B	124	 2% 85% 7% 5%
1	C	124	 3% 95% 5%
1	D	124	 2% 89% 7% 5%
1	E	124	 8% 90% 6% 5%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	3	
2	G	3	

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
4	PG4	B	201	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5500 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 Subtilase cytotoxin subunit B-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	118	Total 948	C 598	N 153	O 189	S 8	0	3	0
1	B	118	Total 945	C 596	N 153	O 188	S 8	0	3	0
1	C	124	Total 985	C 619	N 159	O 199	S 8	0	2	0
1	D	119	Total 942	C 595	N 152	O 186	S 9	0	2	0
1	E	120	Total 947	C 597	N 154	O 188	S 8	0	1	0

There are 20 discrepancies between the modelled and reference sequences:

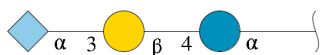
Chain	Residue	Modelled	Actual	Comment	Reference
A	142	LEU	-	expression tag	UNP A0A716TY65
A	143	GLU	-	expression tag	UNP A0A716TY65
A	144	SER	-	expression tag	UNP A0A716TY65
A	145	ASP	-	expression tag	UNP A0A716TY65
B	142	LEU	-	expression tag	UNP A0A716TY65
B	143	GLU	-	expression tag	UNP A0A716TY65
B	144	SER	-	expression tag	UNP A0A716TY65
B	145	ASP	-	expression tag	UNP A0A716TY65
C	142	LEU	-	expression tag	UNP A0A716TY65
C	143	GLU	-	expression tag	UNP A0A716TY65
C	144	SER	-	expression tag	UNP A0A716TY65
C	145	ASP	-	expression tag	UNP A0A716TY65
D	142	LEU	-	expression tag	UNP A0A716TY65
D	143	GLU	-	expression tag	UNP A0A716TY65
D	144	SER	-	expression tag	UNP A0A716TY65
D	145	ASP	-	expression tag	UNP A0A716TY65
E	142	LEU	-	expression tag	UNP A0A716TY65
E	143	GLU	-	expression tag	UNP A0A716TY65
E	144	SER	-	expression tag	UNP A0A716TY65

*Continued on next page...*

Continued from previous page...

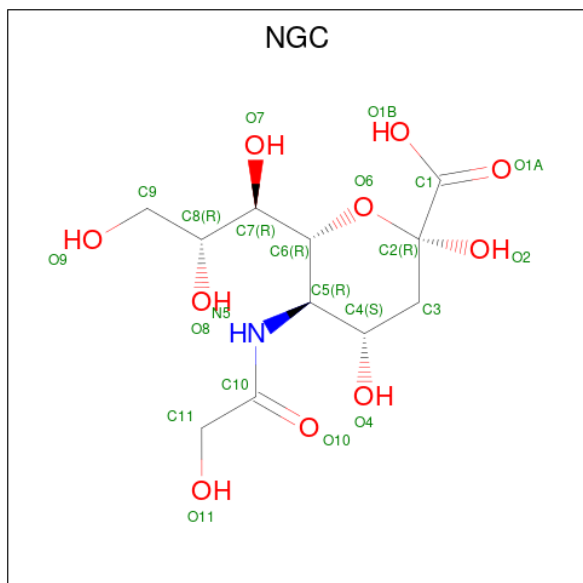
Chain	Residue	Modelled	Actual	Comment	Reference
E	145	ASP	-	expression tag	UNP A0A716TY65

- Molecule 2 is an oligosaccharide called N-glycolyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	F	3	44	23	1	20	0	0	0
2	G	3	44	23	1	20	0	0	0

- Molecule 3 is N-glycolyl-alpha-neuraminic acid (CCD ID: NGC) (formula:  $C_{11}H_{19}NO_{10}$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	22	11	1	10	0	0
3	D	1	22	11	1	10	0	0

- Molecule 4 is TETRAETHYLENE GLYCOL (CCD ID: PG4) (formula:  $C_8H_{18}O_5$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total C O 13 8 5	0	0
4	B	1	Total C O 13 8 5	0	0

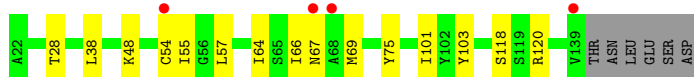
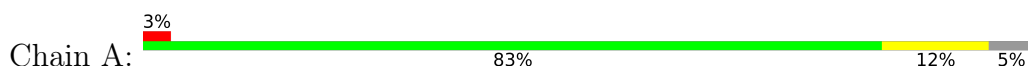
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	115	Total O 115 115	0	0
5	B	120	Total O 120 120	0	0
5	C	104	Total O 104 104	0	0
5	D	121	Total O 121 121	0	0
5	E	115	Total O 115 115	0	0

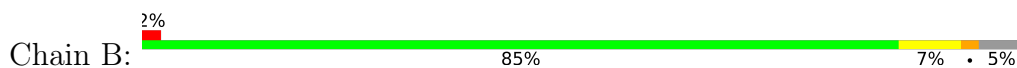
### 3 Residue-property plots [i](#)

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.

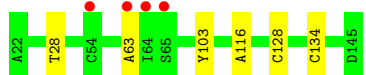
- Molecule 1: Subtilase cytotoxin subunit B-like protein



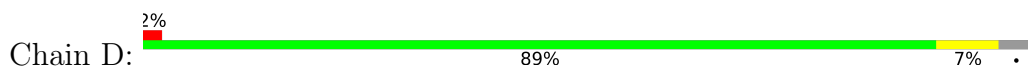
- Molecule 1: Subtilase cytotoxin subunit B-like protein



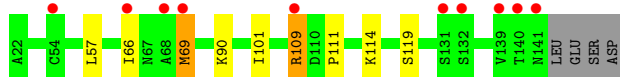
- Molecule 1: Subtilase cytotoxin subunit B-like protein



- Molecule 1: Subtilase cytotoxin subunit B-like protein



- Molecule 1: Subtilase cytotoxin subunit B-like protein




- Molecule 2: N-glycolyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-alpha-D-glucopyranose

Chain F:  33% 67%

GLC1  
GAL2  
MGC3

- Molecule 2: N-glycolyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-alpha-D-glucopyranose

Chain G:  67% 33%

GLC1  
GAL2  
MGC3

## 4 Data and refinement statistics i

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	189.09Å 189.09Å 41.21Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	27.76 – 1.24 27.76 – 1.24	Depositor EDS
% Data completeness (in resolution range)	99.8 (27.76-1.24) 99.7 (27.76-1.24)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.01 (at 1.24Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.138 , 0.160 0.139 , 0.161	Depositor DCC
$R_{free}$ test set	7789 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.3	Xtrriage
Anisotropy	0.531	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 43.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.014 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	5500	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.67% 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: GAL, PG4, NGC, GLC

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	0.87	1/969 (0.1%)	0.99	3/1311 (0.2%)
1	B	1.00	2/966 (0.2%)	1.00	3/1307 (0.2%)
1	C	0.88	0/1006	0.97	2/1362 (0.1%)
1	D	0.71	0/966	0.84	1/1307 (0.1%)
1	E	0.87	2/968 (0.2%)	0.88	0/1310
All	All	0.87	5/4875 (0.1%)	0.94	9/6597 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	110	ASP	C-N	11.57	1.48	1.34
1	E	69	MET	SD-CE	-7.88	1.59	1.79
1	E	111	PRO	C-O	-7.73	1.14	1.24
1	B	59	HIS	CE1-NE2	-6.23	1.26	1.32
1	A	118	SER	C-O	-5.53	1.17	1.23

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	110	ASP	O-C-N	-9.61	112.24	121.18
1	B	110	ASP	CA-C-N	7.75	127.40	119.19
1	B	110	ASP	C-N-CA	7.75	127.40	119.19
1	C	63	ALA	CA-C-N	-5.92	114.54	122.77
1	C	63	ALA	C-N-CA	-5.92	114.54	122.77
1	A	66	ILE	CA-C-N	-5.82	113.40	122.37
1	A	66	ILE	C-N-CA	-5.82	113.40	122.37
1	D	118	SER	N-CA-CB	-5.31	103.29	111.46
1	A	118	SER	CA-C-O	-5.19	115.68	121.23

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	948	0	889	17	0
1	B	945	0	888	15	0
1	C	985	0	923	4	0
1	D	942	0	890	5	0
1	E	947	0	891	12	0
2	F	44	0	28	1	0
2	G	44	0	28	3	0
3	A	22	0	7	0	0
3	D	22	0	8	0	0
4	B	26	0	36	12	0
5	A	115	0	0	2	0
5	B	120	0	0	1	0
5	C	104	0	0	0	0
5	D	121	0	0	1	0
5	E	115	0	0	1	0
All	All	5500	0	4588	54	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 6.

All (54) 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:69:MET:HE3	1:A:101:ILE:HG21	1.20	1.07
1:E:57:LEU:HD11	1:E:69:MET:HE1	1.47	0.95
1:A:55:ILE:HD11	1:A:69:MET:HE2	1.60	0.83
1:A:64:ILE:O	1:A:67:ASN:HB3	1.79	0.82
1:A:57:LEU:HD11	1:A:69:MET:HE1	1.63	0.79
1:E:69:MET:HE3	1:E:101:ILE:HG21	1.64	0.78
1:B:79:LYS:HZ1	4:B:202:PG4:H22	1.51	0.75
1:E:114:LYS:HD2	1:E:119:SER:HB3	1.68	0.74
1:E:57:LEU:HD11	1:E:69:MET:CE	2.20	0.71

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:57:LEU:CD1	1:E:69:MET:HE1	2.19	0.70
1:A:55:ILE:CD1	1:A:69:MET:HE2	2.21	0.70
1:A:67:ASN:ND2	5:A:301:HOH:O	2.24	0.68
1:B:79:LYS:NZ	4:B:202:PG4:H22	2.11	0.65
1:A:57:LEU:HD11	1:A:69:MET:CE	2.29	0.61
1:A:55:ILE:HD11	1:A:69:MET:CE	2.29	0.60
1:E:114:LYS:CD	1:E:119:SER:HB3	2.31	0.60
1:A:55:ILE:CG1	1:A:69:MET:HE2	2.32	0.59
1:A:69:MET:CE	1:A:101:ILE:HG21	2.13	0.59
1:D:48:LYS:NZ	5:D:301:HOH:O	2.35	0.59
1:B:110:ASP:HB3	4:B:201:PG4:H61	1.84	0.59
1:C:28[A]:THR:HG22	1:C:103:TYR:CZ	2.40	0.56
1:A:69:MET:HE3	1:A:101:ILE:CG2	2.14	0.55
1:E:109:ARG:CD	2:G:2:GAL:O2	2.55	0.55
1:A:48:LYS:HD3	1:A:75:TYR:OH	2.08	0.54
1:B:75:TYR:HA	4:B:202:PG4:H21	1.89	0.54
1:D:54[B]:CYS:SG	1:D:70:CYS:SG	3.02	0.52
1:A:57:LEU:CD1	1:A:69:MET:HE1	2.36	0.52
1:B:110:ASP:CB	4:B:201:PG4:H61	2.40	0.51
1:A:120:ARG:NH1	5:A:304:HOH:O	2.36	0.50
1:A:38[A]:LEU:HD12	1:A:54:CYS:O	2.12	0.49
1:E:109:ARG:HD3	2:G:2:GAL:O2	2.12	0.49
1:C:116:ALA:HB1	1:D:137:PRO:HD3	1.95	0.49
1:B:128:CYS:HA	1:B:134:CYS:HA	1.95	0.48
1:B:28[A]:THR:HG22	1:B:103:TYR:CZ	2.49	0.48
1:B:100:ARG:CZ	1:B:134:CYS:HB2	2.44	0.48
1:D:69:MET:SD	1:D:101:ILE:HG21	2.54	0.48
1:E:90:LYS:HE3	5:E:244:HOH:O	2.16	0.46
1:B:79:LYS:HZ2	4:B:202:PG4:H81	1.81	0.46
4:B:201:PG4:O2	4:B:202:PG4:O1	2.31	0.46
1:A:55:ILE:HG13	1:A:69:MET:HE2	1.98	0.46
1:B:59:HIS:CE1	5:B:302:HOH:O	2.69	0.45
1:B:107:VAL:HG13	4:B:201:PG4:H11	1.97	0.45
1:B:110:ASP:H	4:B:201:PG4:H61	1.81	0.44
1:B:79:LYS:NZ	4:B:202:PG4:H81	2.33	0.44
1:B:76:GLY:HA2	4:B:201:PG4:H12	2.00	0.43
1:B:109:ARG:H	4:B:201:PG4:H31	1.84	0.42
1:E:109:ARG:HD2	2:G:2:GAL:O2	2.20	0.41
1:A:28[A]:THR:HG22	1:A:103:TYR:CZ	2.55	0.41
1:C:28[A]:THR:CG2	1:C:103:TYR:CZ	3.04	0.41
1:E:66:ILE:HD13	1:E:66:ILE:HG21	1.88	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:69:MET:HE3	1:E:101:ILE:CG2	2.44	0.41
1:C:128:CYS:HA	1:C:134:CYS:HA	2.03	0.40
1:D:128:CYS:HA	1:D:134:CYS:HA	2.03	0.40
2:F:2:GAL:H4	2:F:3:NGC:C1	2.52	0.40

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	119/124 (96%)	116 (98%)	3 (2%)	0	100	100
1	B	119/124 (96%)	115 (97%)	4 (3%)	0	100	100
1	C	124/124 (100%)	123 (99%)	1 (1%)	0	100	100
1	D	119/124 (96%)	118 (99%)	1 (1%)	0	100	100
1	E	119/124 (96%)	115 (97%)	4 (3%)	0	100	100
All	All	600/620 (97%)	587 (98%)	13 (2%)	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	104/107 (97%)	104 (100%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	104/107 (97%)	103 (99%)	1 (1%)	68	36
1	C	109/107 (102%)	109 (100%)	0	100	100
1	D	104/107 (97%)	104 (100%)	0	100	100
1	E	104/107 (97%)	103 (99%)	1 (1%)	68	36
All	All	525/535 (98%)	523 (100%)	2 (0%)	84	63

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	109	ARG
1	E	109	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	37	ASN
1	A	77	ASN
1	B	37	ASN
1	C	32	ASN
1	C	37	ASN
1	C	77	ASN
1	C	84	ASN
1	D	37	ASN
1	D	77	ASN
1	E	37	ASN
1	E	51	GLN
1	E	77	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

6 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	GLC	F	1	2	12,12,12	1.26	1 (8%)	17,17,17	1.30	3 (17%)
2	GAL	F	2	2	11,11,12	1.39	2 (18%)	15,15,17	2.56	5 (33%)
2	NGC	F	3	2	21,21,22	1.82	6 (28%)	22,29,32	1.77	7 (31%)
2	GLC	G	1	2	12,12,12	1.25	1 (8%)	17,17,17	1.93	6 (35%)
2	GAL	G	2	2	11,11,12	1.98	1 (9%)	15,15,17	2.51	8 (53%)
2	NGC	G	3	2	21,21,22	2.03	9 (42%)	22,29,32	2.03	5 (22%)

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	GLC	F	1	2	-	0/2/22/22	0/1/1/1
2	GAL	F	2	2	-	0/2/19/22	0/1/1/1
2	NGC	F	3	2	-	0/20/36/40	0/1/1/1
2	GLC	G	1	2	-	1/2/22/22	0/1/1/1
2	GAL	G	2	2	-	0/2/19/22	0/1/1/1
2	NGC	G	3	2	-	0/20/36/40	0/1/1/1

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	2	GAL	O5-C1	5.77	1.53	1.43
2	F	3	NGC	C10-N5	5.19	1.45	1.34
2	G	3	NGC	C11-C10	4.36	1.54	1.51
2	G	3	NGC	C10-N5	3.52	1.41	1.34
2	G	3	NGC	O10-C10	-3.43	1.16	1.23
2	F	2	GAL	O5-C1	3.32	1.49	1.43
2	G	1	GLC	O5-C1	3.20	1.50	1.42

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	1	GLC	O5-C1	3.09	1.50	1.42
2	G	3	NGC	O6-C6	2.84	1.48	1.44
2	F	3	NGC	C3-C4	-2.55	1.47	1.52
2	F	3	NGC	O6-C2	2.55	1.48	1.43
2	G	3	NGC	C6-C5	-2.51	1.49	1.53
2	F	3	NGC	O4-C4	2.45	1.48	1.43
2	G	3	NGC	O1A-C1	2.43	1.29	1.22
2	F	3	NGC	O1A-C1	2.41	1.29	1.22
2	F	2	GAL	O3-C3	2.32	1.48	1.43
2	G	3	NGC	C5-N5	2.26	1.49	1.45
2	G	3	NGC	O6-C2	2.16	1.47	1.43
2	G	3	NGC	O7-C7	-2.16	1.37	1.43
2	F	3	NGC	C5-N5	2.05	1.49	1.45

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	2	GAL	C1-C2-C3	-5.98	100.94	109.64
2	G	2	GAL	C1-O5-C5	-5.16	105.27	112.19
2	G	3	NGC	O10-C10-C11	5.13	122.95	119.62
2	F	2	GAL	C1-O5-C5	-5.11	105.34	112.19
2	G	2	GAL	O5-C1-C2	-4.84	99.23	110.79
2	G	3	NGC	O1B-C1-C2	4.26	123.79	112.71
2	F	3	NGC	O10-C10-C11	3.36	121.80	119.62
2	G	2	GAL	O2-C2-C3	-3.23	103.46	110.15
2	G	3	NGC	O4-C4-C3	-3.19	101.94	109.86
2	G	1	GLC	O6-C6-C5	-3.15	100.62	111.33
2	F	3	NGC	O1A-C1-C2	-3.11	116.12	122.85
2	F	2	GAL	O2-C2-C3	-3.11	103.71	110.15
2	F	2	GAL	O2-C2-C1	-3.08	102.16	109.22
2	G	3	NGC	O1A-C1-C2	-3.06	116.23	122.85
2	F	2	GAL	O5-C1-C2	-3.06	103.50	110.79
2	G	1	GLC	O3-C3-C4	3.02	117.50	110.38
2	F	3	NGC	C6-C5-N5	-2.99	106.13	110.91
2	G	1	GLC	O4-C4-C5	-2.94	102.08	109.32
2	G	2	GAL	O3-C3-C2	2.92	116.01	110.05
2	F	3	NGC	O1B-C1-C2	2.89	120.24	112.71
2	G	2	GAL	C1-C2-C3	-2.58	105.89	109.64
2	F	3	NGC	O11-C11-C10	-2.55	108.78	113.95
2	F	1	GLC	O2-C2-C3	-2.54	104.40	110.38
2	G	2	GAL	O2-C2-C1	-2.49	103.53	109.22
2	G	2	GAL	O3-C3-C4	2.37	115.96	110.38

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	1	GLC	O1-C1-O5	2.31	117.28	110.41
2	G	2	GAL	C2-C3-C4	2.26	114.84	110.86
2	F	3	NGC	O10-C10-N5	-2.25	119.15	122.95
2	G	1	GLC	O5-C5-C4	2.23	113.72	109.70
2	F	1	GLC	O5-C5-C6	-2.21	100.96	106.44
2	G	1	GLC	O5-C5-C6	2.18	111.83	106.44
2	G	3	NGC	O6-C2-C1	2.10	111.67	107.72
2	F	1	GLC	C1-C2-C3	-2.09	106.10	110.36
2	F	3	NGC	O6-C2-C1	2.08	111.64	107.72

There are no chirality outliers.

All (1) torsion outliers are listed below:

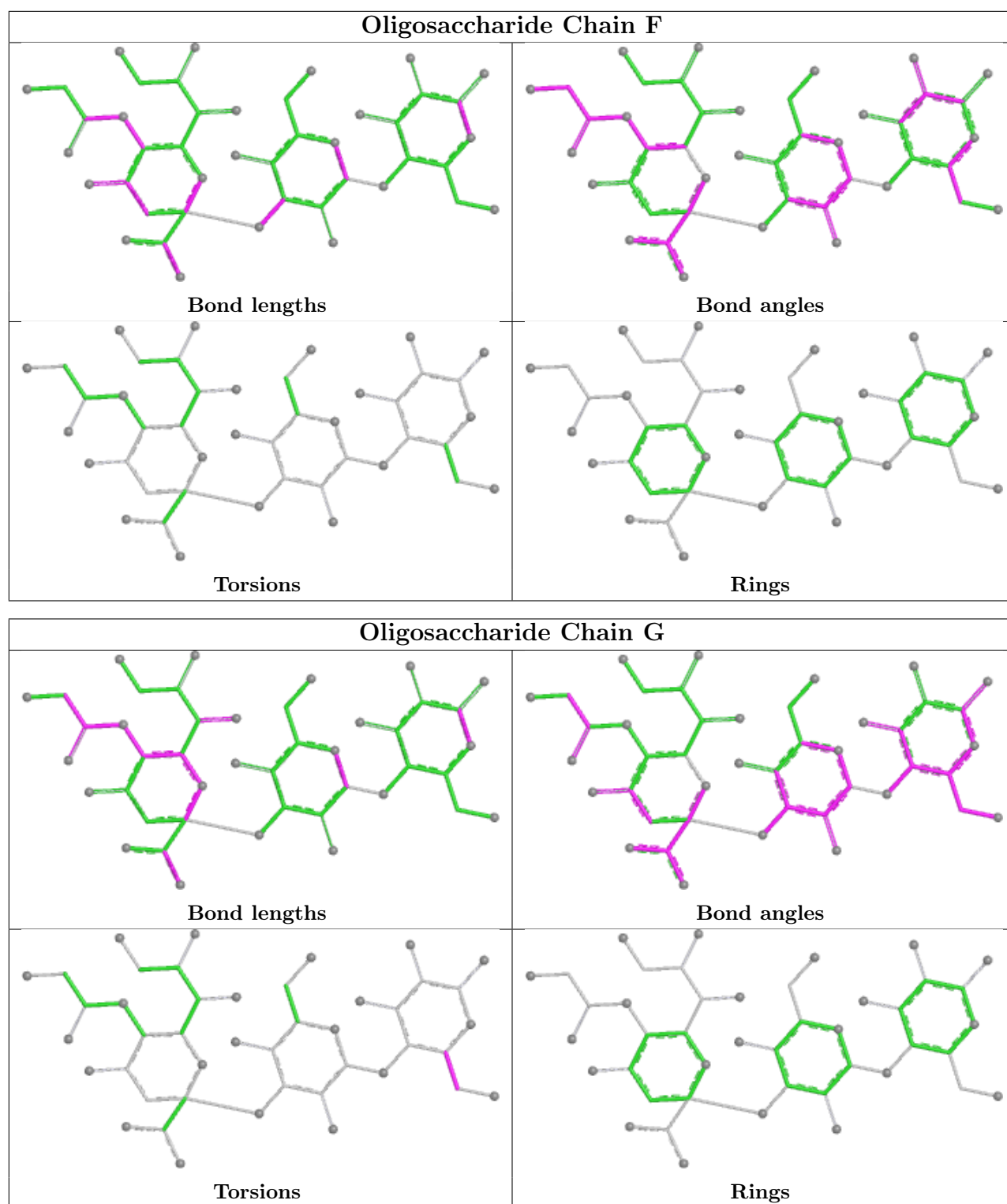
Mol	Chain	Res	Type	Atoms
2	G	1	GLC	O5-C5-C6-O6

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	2	GAL	1	0
2	F	3	NGC	1	0
2	G	2	GAL	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](#)

4 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
3	NGC	D	201	-	22,22,22	2.69	6 (27%)	25,32,32	2.27	8 (32%)
4	PG4	B	202	-	12,12,12	0.55	0	11,11,11	0.60	0
4	PG4	B	201	-	12,12,12	0.67	0	11,11,11	1.01	1 (9%)
3	NGC	A	201	-	22,22,22	1.96	9 (40%)	25,32,32	1.92	7 (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	NGC	D	201	-	-	2/22/40/40	0/1/1/1
4	PG4	B	202	-	-	5/10/10/10	-
4	PG4	B	201	-	-	5/10/10/10	-
3	NGC	A	201	-	-	3/22/40/40	0/1/1/1

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	201	NGC	C11-C10	6.44	1.56	1.51
3	D	201	NGC	O6-C2	5.83	1.49	1.43
3	D	201	NGC	O10-C10	-5.73	1.11	1.23
3	A	201	NGC	O6-C2	4.18	1.47	1.43
3	D	201	NGC	O6-C6	3.96	1.49	1.44
3	A	201	NGC	C10-N5	3.52	1.41	1.34
3	A	201	NGC	O10-C10	-3.35	1.16	1.23
3	D	201	NGC	C5-N5	3.02	1.50	1.45
3	D	201	NGC	C10-N5	3.00	1.40	1.34
3	A	201	NGC	O7-C7	-2.57	1.36	1.43
3	A	201	NGC	C6-C5	-2.43	1.49	1.53
3	A	201	NGC	O6-C6	2.40	1.47	1.44
3	A	201	NGC	C11-C10	2.26	1.53	1.51
3	A	201	NGC	C3-C4	-2.22	1.49	1.53
3	A	201	NGC	C5-N5	2.19	1.49	1.45

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	201	NGC	O10-C10-C11	-5.06	116.34	119.62
3	D	201	NGC	O10-C10-N5	4.49	130.56	122.95
3	D	201	NGC	C5-N5-C10	-4.22	116.68	122.90
3	A	201	NGC	O6-C6-C5	-3.97	106.21	109.84
3	D	201	NGC	O6-C6-C5	-3.96	106.22	109.84
3	D	201	NGC	O6-C6-C7	-3.56	101.10	106.65
3	A	201	NGC	C5-N5-C10	-3.45	117.81	122.90
3	A	201	NGC	O1A-C1-C2	-3.41	118.16	123.85
3	A	201	NGC	O10-C10-N5	3.11	128.21	122.95
3	D	201	NGC	O4-C4-C5	2.86	116.33	109.84
3	A	201	NGC	O4-C4-C3	-2.77	103.52	109.97
3	D	201	NGC	O4-C4-C3	-2.75	103.58	109.97
3	A	201	NGC	C4-C5-N5	-2.17	106.16	110.44
4	B	201	PG4	O4-C6-C5	-2.14	100.58	110.35
3	D	201	NGC	C11-C10-N5	-2.05	112.93	116.42
3	A	201	NGC	O1B-C1-O1A	-2.05	117.29	123.86

There are no chirality outliers.

All (15) torsion outliers are listed below:

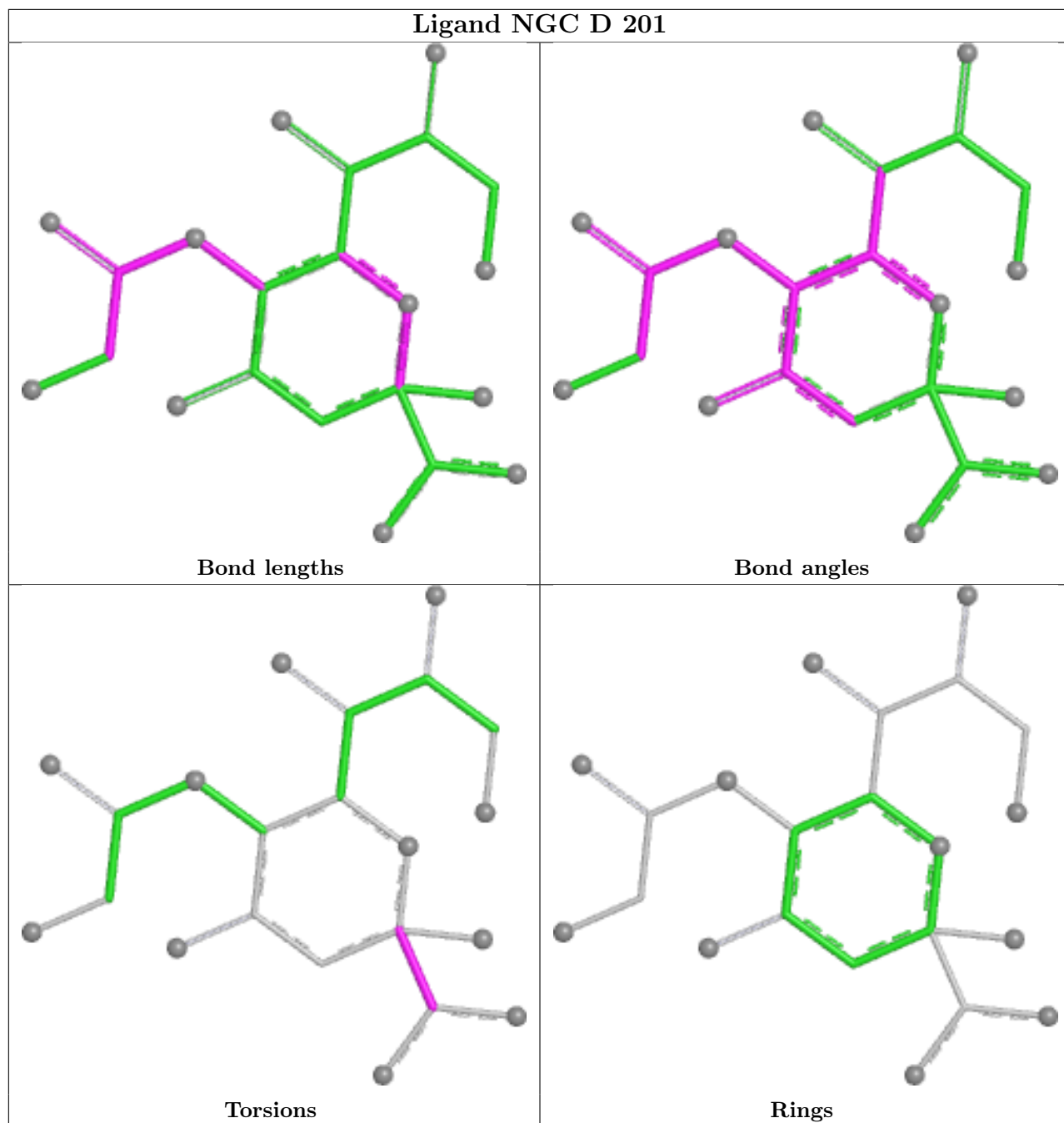
Mol	Chain	Res	Type	Atoms
4	B	201	PG4	O4-C7-C8-O5
4	B	202	PG4	O1-C1-C2-O2
4	B	202	PG4	O2-C3-C4-O3
4	B	201	PG4	C8-C7-O4-C6
4	B	202	PG4	C4-C3-O2-C2
4	B	202	PG4	C8-C7-O4-C6
4	B	201	PG4	O3-C5-C6-O4
3	A	201	NGC	O7-C7-C8-C9
4	B	201	PG4	C6-C5-O3-C4
4	B	201	PG4	C5-C6-O4-C7
3	A	201	NGC	O1B-C1-C2-O6
3	D	201	NGC	O1B-C1-C2-O6
4	B	202	PG4	C3-C4-O3-C5
3	A	201	NGC	O1B-C1-C2-C3
3	D	201	NGC	O1B-C1-C2-C3

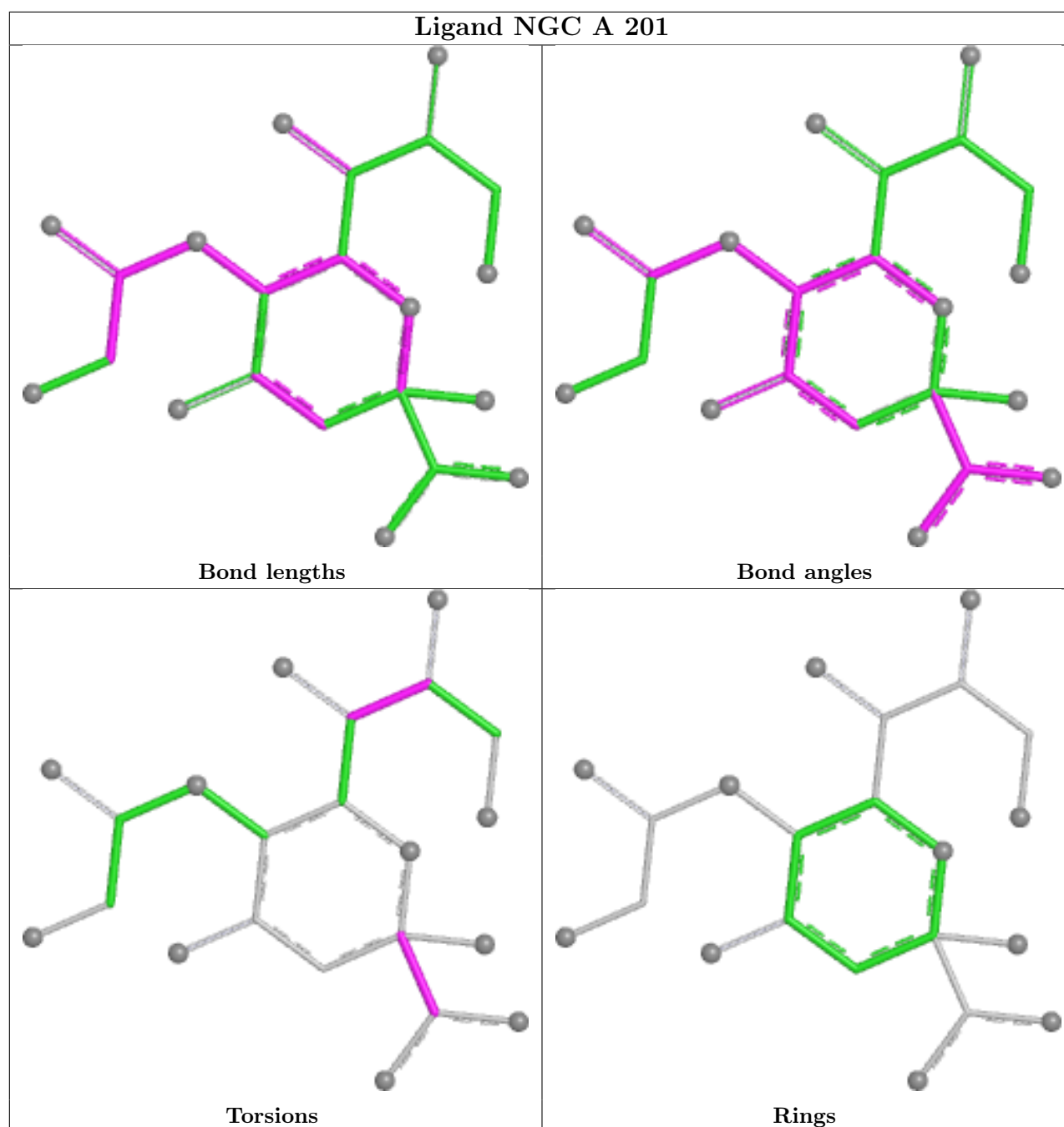
There are no ring outliers.

2 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	202	PG4	6	0
4	B	201	PG4	7	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 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	A	118/124 (95%)	-0.19	4 (3%) 48 48	6, 13, 26, 33	3 (2%)
1	B	118/124 (95%)	-0.17	2 (1%) 69 68	6, 15, 25, 44	3 (2%)
1	C	124/124 (100%)	-0.06	4 (3%) 50 51	6, 16, 31, 47	2 (1%)
1	D	119/124 (95%)	-0.08	2 (1%) 69 68	6, 15, 28, 47	2 (1%)
1	E	120/124 (96%)	0.02	10 (8%) 17 17	6, 16, 29, 50	1 (0%)
All	All	599/620 (96%)	-0.10	22 (3%) 45 45	6, 15, 28, 50	11 (1%)

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	109	ARG	3.8
1	A	67	ASN	3.6
1	A	54	CYS	3.5
1	B	139	VAL	3.3
1	E	141	ASN	3.1
1	A	68	ALA	2.9
1	C	64	ILE	2.8
1	C	63	ALA	2.6
1	D	140	THR	2.4
1	E	66	ILE	2.4
1	C	54	CYS	2.4
1	A	139	VAL	2.3
1	E	132	SER	2.2
1	E	139	VAL	2.2
1	C	65	SER	2.2
1	E	140	THR	2.1
1	E	54	CYS	2.1
1	E	68	ALA	2.1
1	E	69	MET	2.1
1	D	132	SER	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	54	CYS	2.1
1	E	131	SER	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](#)

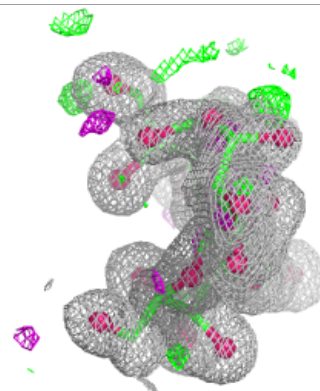
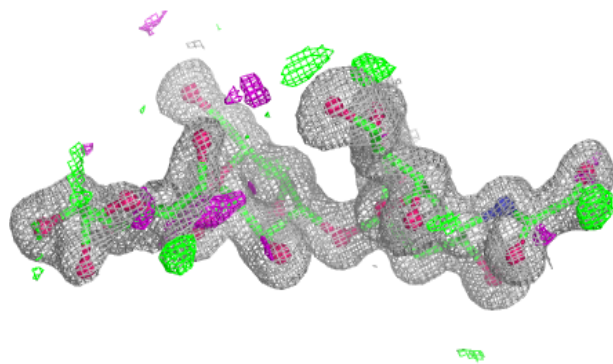
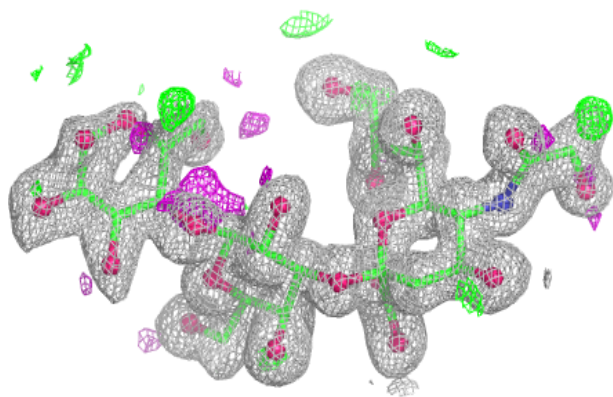
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	GLC	G	1	12/12	0.71	0.14	33,37,50,52	0
2	GLC	F	1	12/12	0.83	0.13	20,36,43,47	0
2	GAL	G	2	11/12	0.84	0.11	26,30,36,38	0
2	NGC	G	3	21/22	0.93	0.09	18,21,28,33	0
2	GAL	F	2	11/12	0.95	0.07	16,19,21,22	0
2	NGC	F	3	21/22	0.96	0.06	14,15,21,23	0

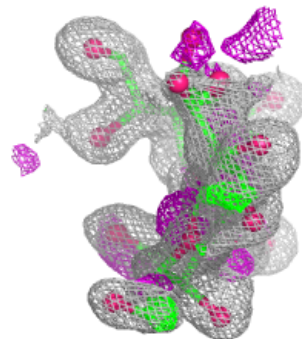
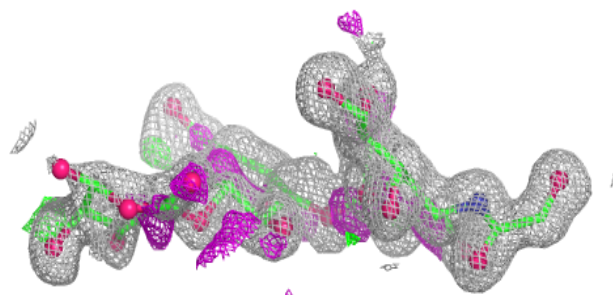
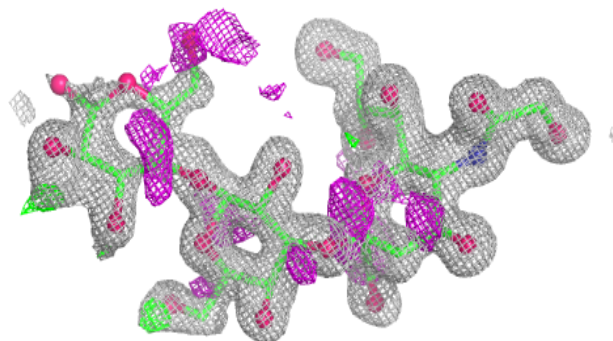
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain F:**

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

**Electron density around Chain G:**

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



## 6.4 Ligands

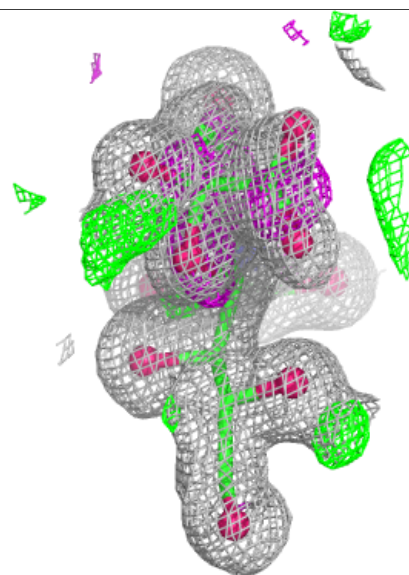
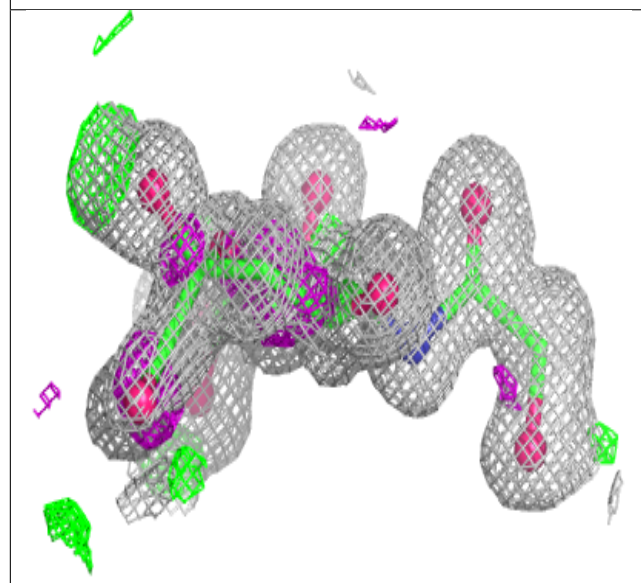
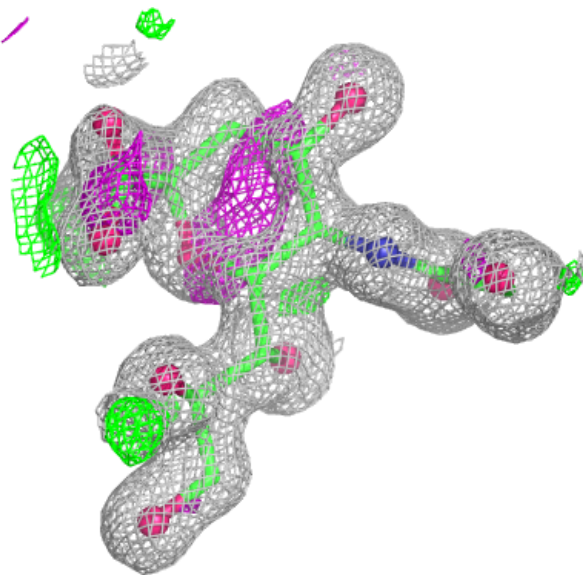
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

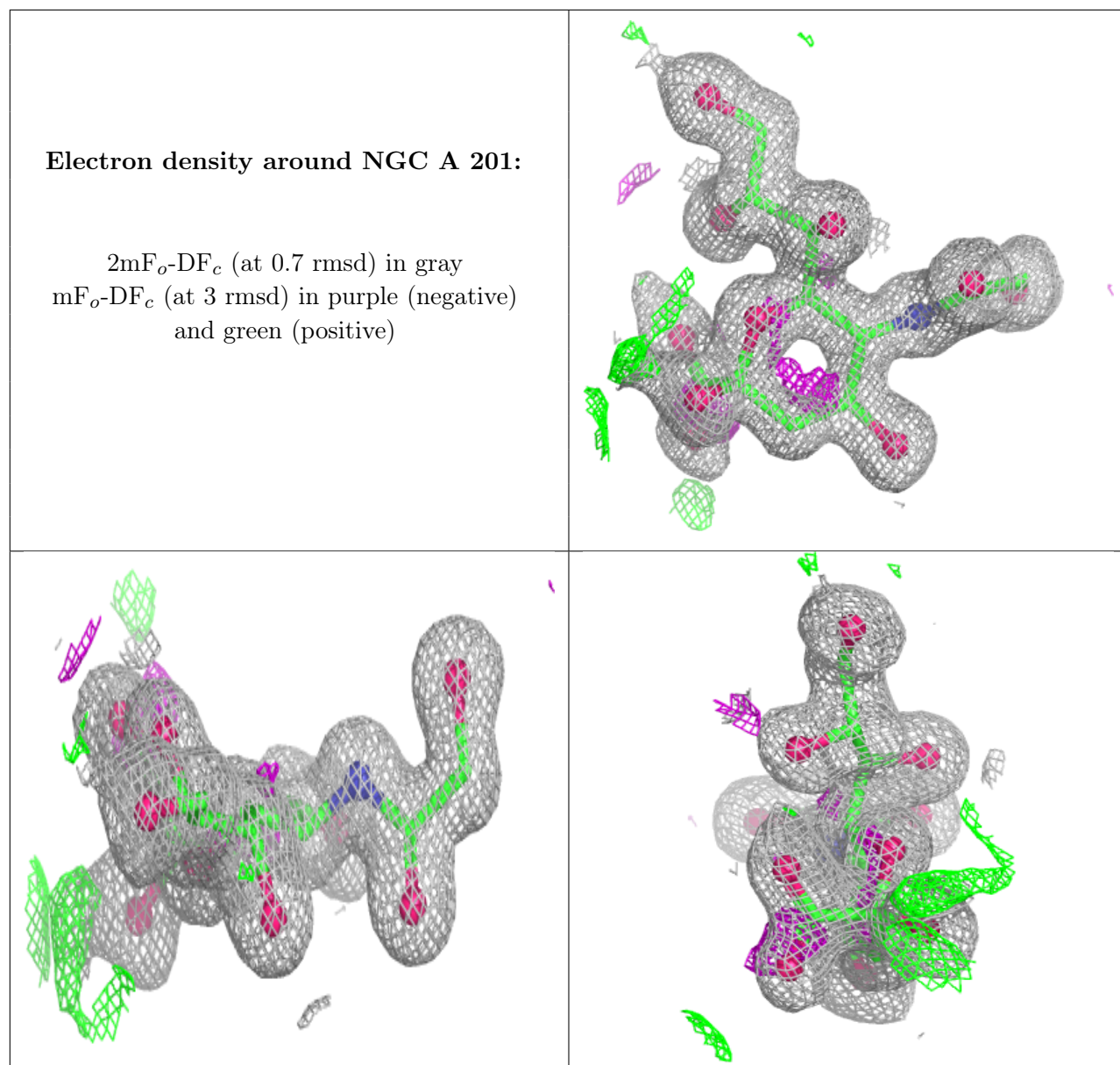
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	PG4	B	202	13/13	0.78	0.16	35,41,52,54	0
4	PG4	B	201	13/13	0.85	0.15	13,28,41,52	0
3	NGC	D	201	22/22	0.94	0.09	12,18,27,45	0
3	NGC	A	201	22/22	0.95	0.07	15,18,23,28	0

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 NGC D 201:**

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





## 6.5 Other polymers ⓘ

There are no such residues in this entry.