



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 10, 2020 – 11:27 AM BST

PDB ID : 4ZH1  
Title : Complement factor H in complex with the GM1 glycan  
Authors : Blaum, B.S.; Stehle, T.  
Deposited on : 2015-04-24  
Resolution : 2.24 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 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 : 2.13.1  
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) : 2.13.1

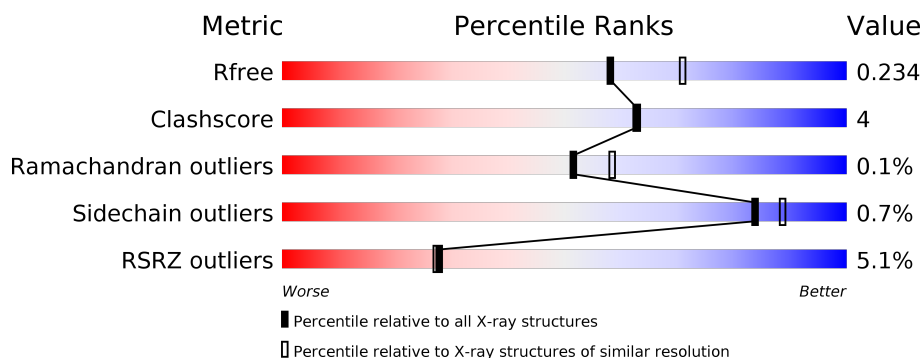
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.24 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.



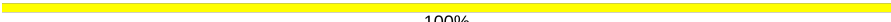

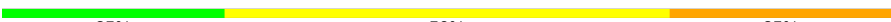
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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 on 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	317	<div> <div style="width: 85%;"></div> <div style="width: 8%;"></div> <div style="width: 8%;"></div> </div>
1	B	317	<div> <div style="width: 80%;"></div> <div style="width: 11%;"></div> <div style="width: 8%;"></div> </div>
1	C	317	<div> <div style="width: 4%;"></div> <div style="width: 85%;"></div> <div style="width: 6%;"></div> <div style="width: 8%;"></div> </div>
2	D	129	<div> <div style="width: 3%;"></div> <div style="width: 85%;"></div> <div style="width: 12%;"></div> <div style="width: 5%;"></div> </div>
2	E	129	<div> <div style="width: 2%;"></div> <div style="width: 88%;"></div> <div style="width: 6%;"></div> <div style="width: 5%;"></div> </div>
2	F	129	<div> <div style="width: 29%;"></div> <div style="width: 70%;"></div> <div style="width: 10%;"></div> <div style="width: 19%;"></div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	2	 100%
4	H	5	 20%80%
5	I	4	 25%50%25%

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 10322 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 Complement C3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	293	Total	C	N	O	S	0	1	0
			2311	1484	391	427	9			
1	B	292	Total	C	N	O	S	0	1	0
			2295	1473	381	432	9			
1	C	291	Total	C	N	O	S	0	3	0
			2280	1471	378	422	9			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	GLY	-	expression tag	UNP P01024
A	-5	PRO	-	expression tag	UNP P01024
A	-4	LEU	-	expression tag	UNP P01024
A	-3	GLY	-	expression tag	UNP P01024
A	-2	SER	-	expression tag	UNP P01024
A	-1	PRO	-	expression tag	UNP P01024
A	0	GLU	-	expression tag	UNP P01024
A	1	PHE	-	expression tag	UNP P01024
A	2	ARG	-	expression tag	UNP P01024
A	17	ALA	CYS	conflict	UNP P01024
B	-6	GLY	-	expression tag	UNP P01024
B	-5	PRO	-	expression tag	UNP P01024
B	-4	LEU	-	expression tag	UNP P01024
B	-3	GLY	-	expression tag	UNP P01024
B	-2	SER	-	expression tag	UNP P01024
B	-1	PRO	-	expression tag	UNP P01024
B	0	GLU	-	expression tag	UNP P01024
B	1	PHE	-	expression tag	UNP P01024
B	2	ARG	-	expression tag	UNP P01024
B	17	ALA	CYS	conflict	UNP P01024
C	-6	GLY	-	expression tag	UNP P01024
C	-5	PRO	-	expression tag	UNP P01024
C	-4	LEU	-	expression tag	UNP P01024

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP P01024
C	-2	SER	-	expression tag	UNP P01024
C	-1	PRO	-	expression tag	UNP P01024
C	0	GLU	-	expression tag	UNP P01024
C	1	PHE	-	expression tag	UNP P01024
C	2	ARG	-	expression tag	UNP P01024
C	17	ALA	CYS	conflict	UNP P01024

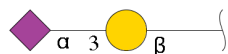
- Molecule 2 is a protein called Complement factor H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	126	Total	C	N	O	S	0	2	0
			1011	637	176	189	9			
2	E	122	Total	C	N	O	S	0	2	0
			989	620	176	184	9			
2	F	104	Total	C	N	O	S	0	1	0
			807	512	135	153	7			

There are 12 discrepancies between the modelled and reference sequences:

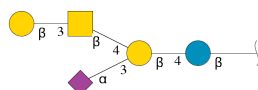
Chain	Residue	Modelled	Actual	Comment	Reference
D	1103	GLU	-	expression tag	UNP P08603
D	1104	ALA	-	expression tag	UNP P08603
D	1105	GLU	-	expression tag	UNP P08603
D	1106	PHE	-	expression tag	UNP P08603
E	1103	GLU	-	expression tag	UNP P08603
E	1104	ALA	-	expression tag	UNP P08603
E	1105	GLU	-	expression tag	UNP P08603
E	1106	PHE	-	expression tag	UNP P08603
F	1103	GLU	-	expression tag	UNP P08603
F	1104	ALA	-	expression tag	UNP P08603
F	1105	GLU	-	expression tag	UNP P08603
F	1106	PHE	-	expression tag	UNP P08603

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose.



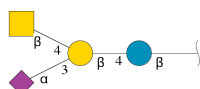
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	G	2	Total	C	N	O	0	0	0
			32	17	1	14			

- Molecule 4 is an oligosaccharide called beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-galactopyranose-(1-4)-[N-acetyl-alpha-neuraminic acid-(2-3)]beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.



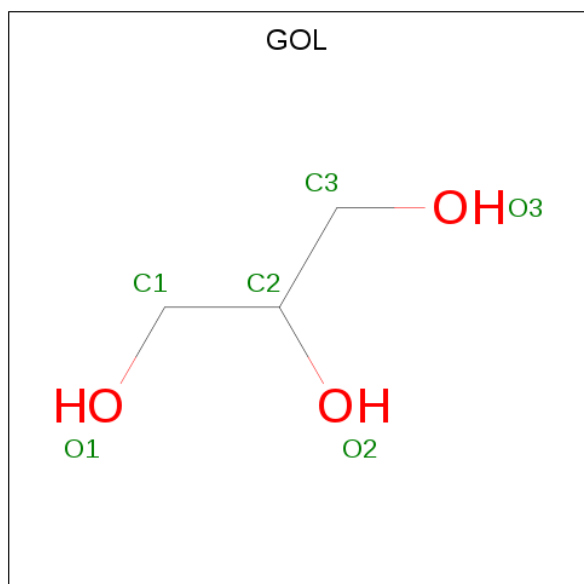
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	H	5	Total	C	N	O	0	0	0
			68	37	2	29			

- Molecule 5 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-[2-acetamido-2-deoxy-beta-D-galactopyranose-(1-4)]beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	4	Total	C	N	O	0	0	0
			57	31	2	24			

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

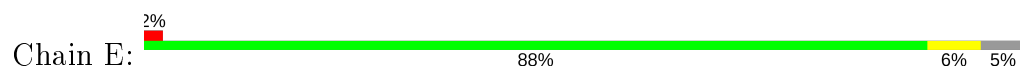
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	143	Total	O	0	0
			143	143		
7	B	101	Total	O	0	0
			101	101		
7	C	73	Total	O	0	0
			73	73		
7	D	75	Total	O	0	0
			75	75		
7	E	50	Total	O	0	0
			50	50		
7	F	6	Total	O	0	0
			6	6		

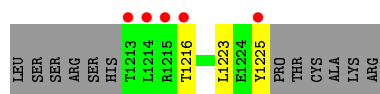
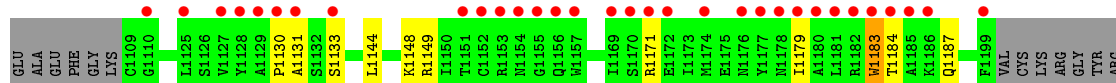




- Molecule 2: Complement factor H



- Molecule 2: Complement factor H



- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose



- Molecule 4: beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-galactopyranose-(1-4)-[N-acetyl-alpha-neuraminic acid-(2-3)]beta-D-galactopyranose-(1-4)-beta-D-glucopyranose



- Molecule 5: N-acetyl-alpha-neuraminic acid-(2-3)-[2-acetamido-2-deoxy-beta-D-galactopyranose-(1-4)]beta-D-galactopyranose-(1-4)-beta-D-glucopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.77Å 82.86Å 86.57Å 112.97° 111.43° 99.63°	Depositor
Resolution (Å)	46.18 – 2.24 46.18 – 2.24	Depositor EDS
% Data completeness (in resolution range)	98.0 (46.18-2.24) 98.0 (46.18-2.24)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.11 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.181 , 0.228 0.188 , 0.234	Depositor DCC
$R_{free}$ test set	3965 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.7	Xtriage
Anisotropy	0.486	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 52.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.009 for -h,-k,h+k+l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10322	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.19% 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: GAL, GOL, SIA, BGC, NGA

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	0.61	0/2360	0.67	0/3198
1	B	0.59	0/2342	0.68	0/3179
1	C	0.51	0/2329	0.63	0/3164
2	D	0.65	0/1038	0.68	0/1408
2	E	0.59	0/1015	0.71	2/1376 (0.1%)
2	F	0.50	0/829	0.67	0/1132
All	All	0.58	0/9913	0.67	2/13457 (0.0%)

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	B	0	1
2	F	0	1
All	All	0	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	1148	LYS	CB-CA-C	-6.51	97.37	110.40
2	E	1148	LYS	CA-CB-CG	5.08	124.57	113.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	143	ASN	Peptide
2	F	1183	TRP	Peptide

## 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	2311	0	2306	21	0
1	B	2295	0	2280	23	0
1	C	2280	0	2256	15	0
2	D	1011	0	967	15	0
2	E	989	0	941	6	0
2	F	807	0	744	7	0
3	G	32	0	28	0	0
4	H	68	0	58	0	0
5	I	57	0	49	1	0
6	A	6	0	8	1	0
6	B	6	0	8	1	0
6	C	6	0	8	3	0
6	D	6	0	8	0	0
7	A	143	0	0	5	0
7	B	101	0	0	3	0
7	C	73	0	0	3	0
7	D	75	0	0	5	0
7	E	50	0	0	0	0
7	F	6	0	0	0	0
All	All	10322	0	9661	84	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 84 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)
2:E:1168:VAL:HG22	2:E:1190[A]:TYR:CE1	1.93	1.03
1:C:121:PRO:O	1:C:178:LYS:HE3	1.69	0.92
2:E:1168:VAL:HG22	2:E:1190[A]:TYR:CZ	2.06	0.90

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:264:ASN:OD1	7:B:501:HOH:O	1.95	0.84
1:A:166:GLU:HG2	7:A:590:HOH:O	1.78	0.82

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	292/317 (92%)	285 (98%)	7 (2%)	0	100	100
1	B	289/317 (91%)	279 (96%)	9 (3%)	1 (0%)	41	44
1	C	292/317 (92%)	285 (98%)	7 (2%)	0	100	100
2	D	126/129 (98%)	124 (98%)	2 (2%)	0	100	100
2	E	120/129 (93%)	119 (99%)	1 (1%)	0	100	100
2	F	101/129 (78%)	97 (96%)	4 (4%)	0	100	100
All	All	1220/1338 (91%)	1189 (98%)	30 (2%)	1 (0%)	51	58

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	303	VAL

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	240/264 (91%)	239 (100%)	1 (0%)	91	93
1	B	243/264 (92%)	240 (99%)	3 (1%)	71	78
1	C	234/264 (89%)	233 (100%)	1 (0%)	91	93
2	D	111/115 (96%)	109 (98%)	2 (2%)	59	66
2	E	109/115 (95%)	109 (100%)	0	100	100
2	F	86/115 (75%)	86 (100%)	0	100	100
All	All	1023/1137 (90%)	1016 (99%)	7 (1%)	84	88

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	170	ASN
2	D	1159	GLU
1	C	222	LYS
1	B	42	GLU
2	D	1144	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	142	ASN
1	B	159	GLN
1	B	266	GLN
1	C	40	GLN
1	C	142	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue 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	GAL	H	4	4	11,11,12	0.57	0	15,15,17	1.13	1 (6%)

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	GAL	H	4	4	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	4	GAL	C1-O5-C5	2.81	116.00	112.19

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

11 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	GAL	G	1	3	12,12,12	0.91	0	17,17,17	1.36	2 (11%)
3	SIA	G	2	3	17,20,21	0.53	0	21,28,31	1.49	4 (19%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	BGC	H	1	4	12,12,12	0.50	0	17,17,17	0.73	0
4	GAL	H	2	4	11,11,12	0.60	0	15,15,17	1.66	4 (26%)
4	NGA	H	3	4	14,14,15	0.57	0	17,19,21	1.66	1 (5%)
4	GAL	H	4	4	11,11,12	0.57	0	15,15,17	1.13	1 (6%)
4	SIA	H	5	4	17,20,21	0.61	0	21,28,31	1.23	2 (9%)
5	BGC	I	1	5	12,12,12	0.86	0	17,17,17	0.96	0
5	GAL	I	2	5	11,11,12	0.43	0	15,15,17	1.69	3 (20%)
5	SIA	I	3	5	17,20,21	0.58	0	21,28,31	1.28	1 (4%)
5	NGA	I	4	5	14,14,15	0.48	0	17,19,21	1.21	1 (5%)

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	GAL	G	1	3	-	1/2/22/22	0/1/1/1
3	SIA	G	2	3	-	0/14/34/38	0/1/1/1
4	BGC	H	1	4	-	0/2/22/22	0/1/1/1
4	GAL	H	2	4	-	2/2/19/22	0/1/1/1
4	NGA	H	3	4	-	2/6/23/26	0/1/1/1
4	GAL	H	4	4	-	0/2/19/22	0/1/1/1
4	SIA	H	5	4	-	0/14/34/38	0/1/1/1
5	BGC	I	1	5	-	2/2/22/22	0/1/1/1
5	GAL	I	2	5	-	1/2/19/22	0/1/1/1
5	SIA	I	3	5	-	6/14/34/38	0/1/1/1
5	NGA	I	4	5	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	3	NGA	O5-C5-C6	5.33	115.56	107.20
3	G	2	SIA	C4-C3-C2	4.10	117.15	109.81
5	I	2	GAL	C1-O5-C5	4.06	117.69	112.19
5	I	3	SIA	C3-C4-C5	-4.03	106.59	111.46
4	H	2	GAL	O3-C3-C2	-3.41	103.46	109.99

There are no chirality outliers.



5 of 14 torsion outliers are listed below:

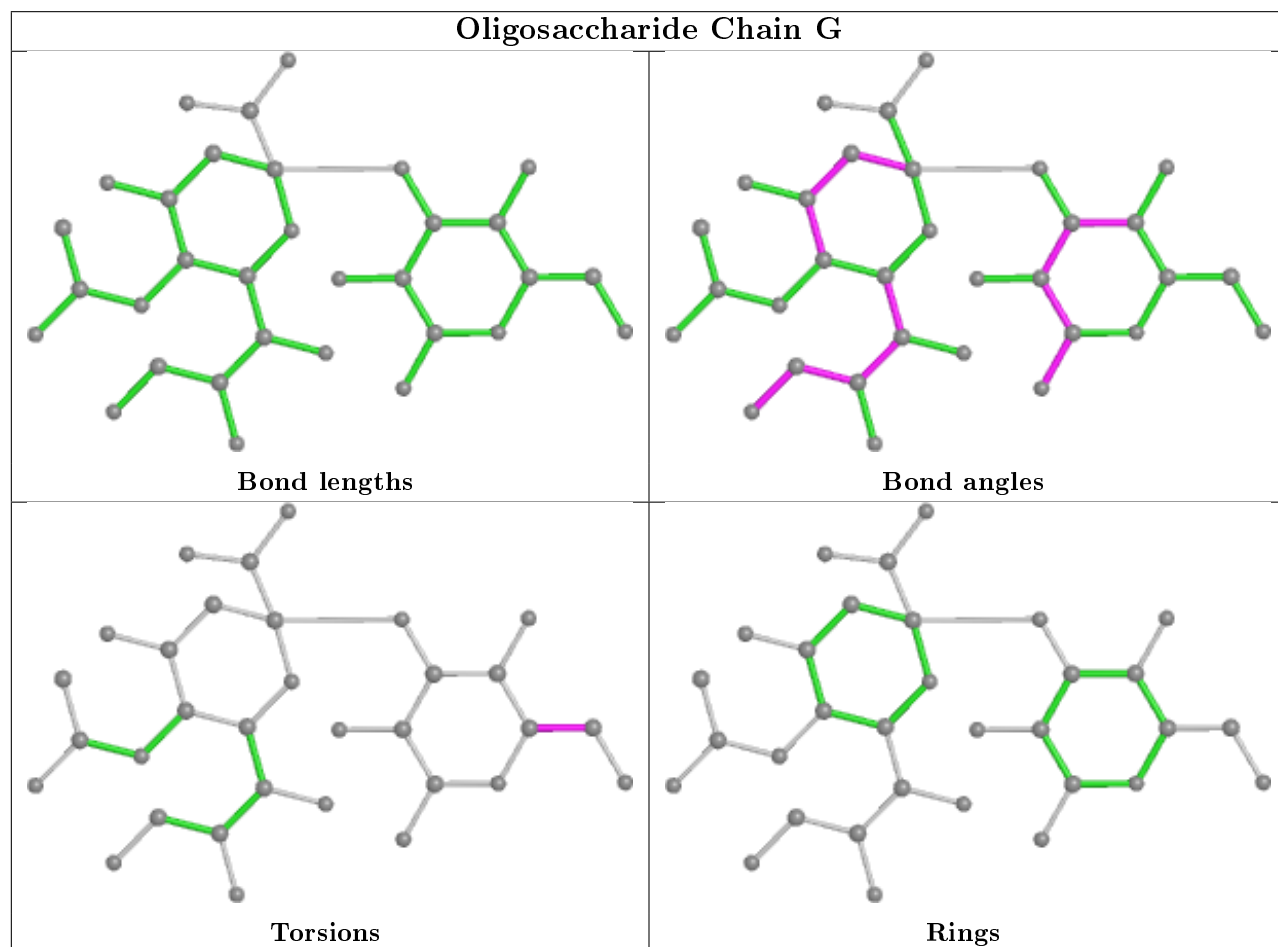
Mol	Chain	Res	Type	Atoms
5	I	3	SIA	O7-C7-C8-C9
5	I	3	SIA	O7-C7-C8-O8
5	I	1	BGC	O5-C5-C6-O6
5	I	1	BGC	C4-C5-C6-O6
4	H	3	NGA	O5-C5-C6-O6

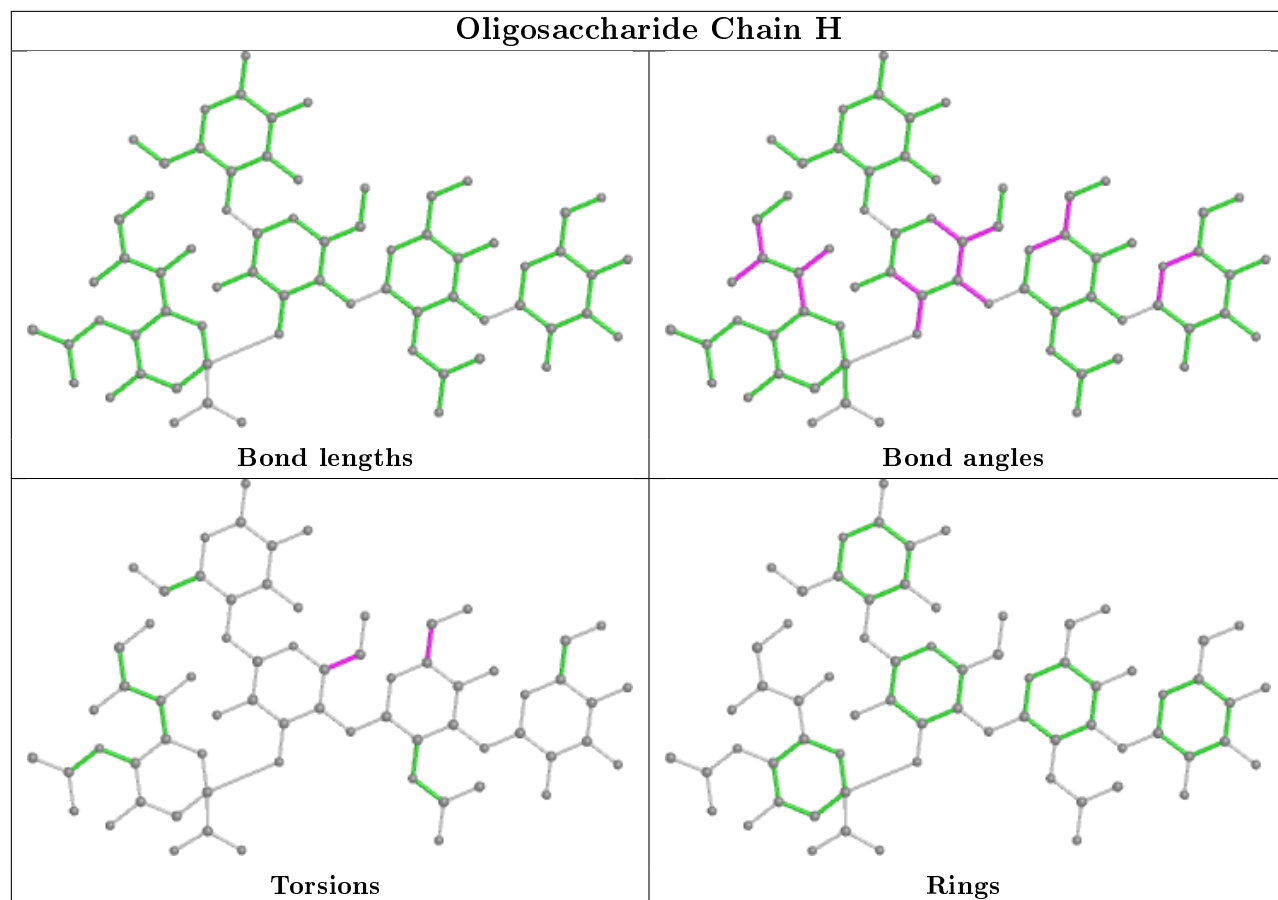
There are no ring outliers.

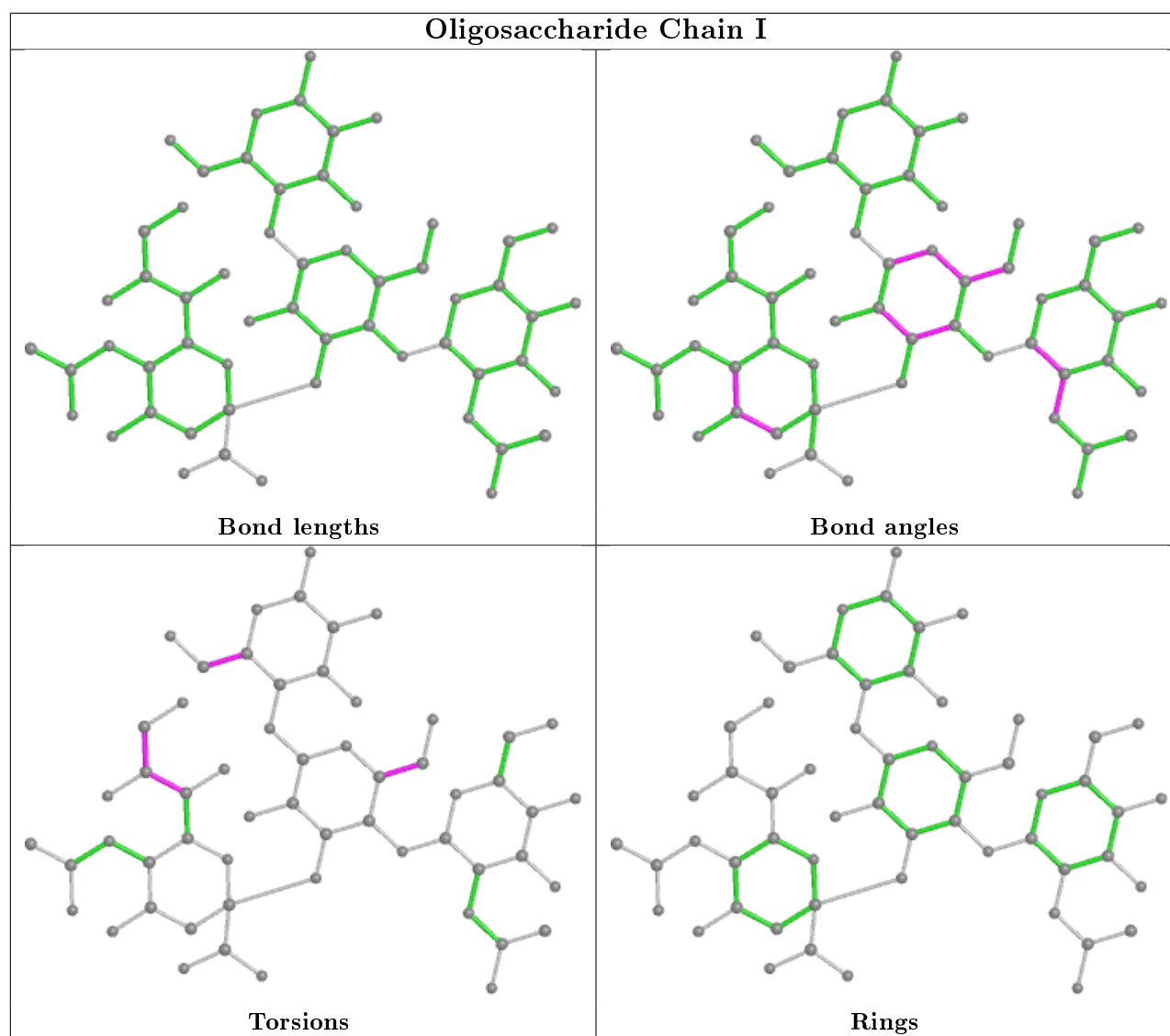
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	I	3	SIA	1	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	$\# Z  > 2$
6	GOL	C	401	-	5,5,5	0.41	0	5,5,5	1.42	1 (20%)
6	GOL	D	1303	-	5,5,5	0.56	0	5,5,5	0.94	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	GOL	B	401	-	5,5,5	0.48	0	5,5,5	0.61	0
6	GOL	A	401	-	5,5,5	0.38	0	5,5,5	0.43	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
6	GOL	C	401	-	-	3/4/4/4	-
6	GOL	D	1303	-	-	1/4/4/4	-
6	GOL	B	401	-	-	4/4/4/4	-
6	GOL	A	401	-	-	2/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	C	401	GOL	C3-C2-C1	-2.62	101.52	111.70

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	401	GOL	O1-C1-C2-O2
6	B	401	GOL	O1-C1-C2-C3
6	C	401	GOL	C1-C2-C3-O3
6	B	401	GOL	C1-C2-C3-O3
6	A	401	GOL	C1-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	401	GOL	3	0
6	B	401	GOL	1	0
6	A	401	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, 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	293/317 (92%)	-0.30	2 (0%) 87 87	37, 54, 85, 124	0
1	B	292/317 (92%)	-0.28	3 (1%) 82 83	39, 58, 94, 130	0
1	C	291/317 (91%)	0.10	14 (4%) 30 30	45, 71, 117, 135	0
2	D	126/129 (97%)	-0.31	4 (3%) 47 47	43, 56, 92, 131	0
2	E	122/129 (94%)	-0.18	3 (2%) 57 58	46, 68, 96, 121	0
2	F	104/129 (80%)	1.55	37 (35%) 0 0	54, 108, 152, 170	0
All	All	1228/1338 (91%)	-0.03	63 (5%) 28 27	37, 63, 117, 170	0

The worst 5 of 63 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	41	TRP	6.3
2	F	1179	ILE	6.0
2	F	1130	PRO	5.9
2	F	1180	ALA	5.8
2	F	1214	LEU	5.6

### 6.2 Non-standard residues in protein, DNA, RNA chains [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.

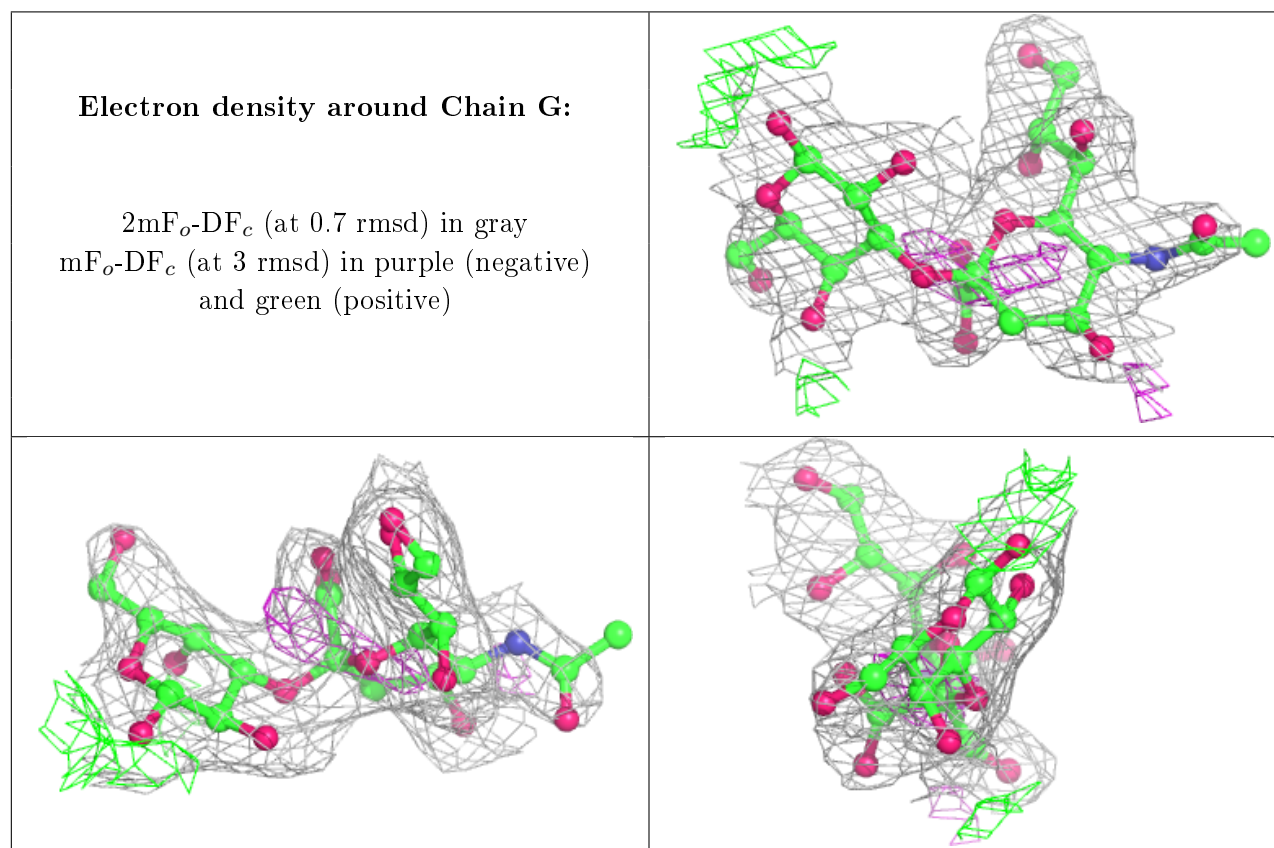
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	GAL	H	4	11/12	0.92	0.10	89,95,103,116	0

## 6.3 Carbohydrates ⓘ

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.

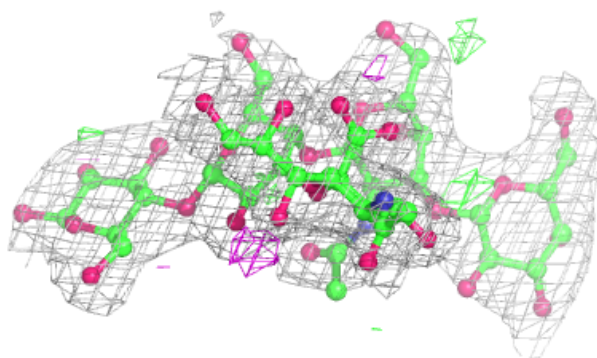
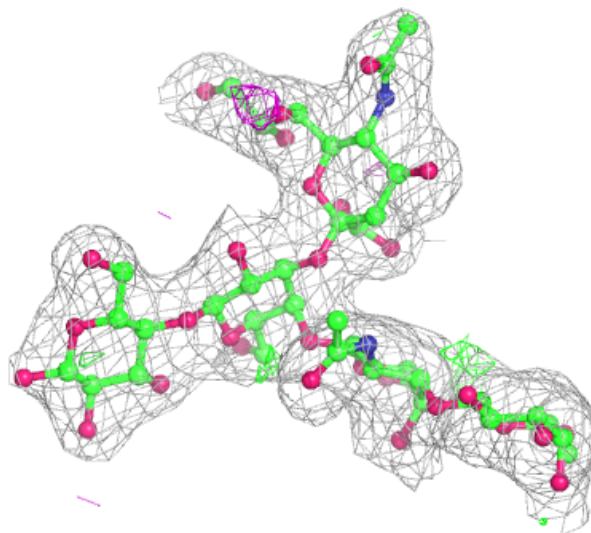
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	BGC	I	1	12/12	0.60	0.33	121,134,140,143	0
3	GAL	G	1	12/12	0.74	0.21	78,98,101,105	0
5	NGA	I	4	14/15	0.90	0.16	97,109,112,120	0
3	SIA	G	2	20/21	0.91	0.25	52,88,104,109	0
5	GAL	I	2	11/12	0.91	0.12	93,107,121,126	0
5	SIA	I	3	20/21	0.92	0.16	93,105,123,125	0
4	GAL	H	4	11/12	0.92	0.10	89,95,103,116	0
4	NGA	H	3	14/15	0.94	0.09	70,83,90,98	0
4	BGC	H	1	12/12	0.95	0.08	66,72,80,87	0
4	GAL	H	2	11/12	0.96	0.06	58,68,78,96	0
4	SIA	H	5	20/21	0.97	0.10	52,59,82,91	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.

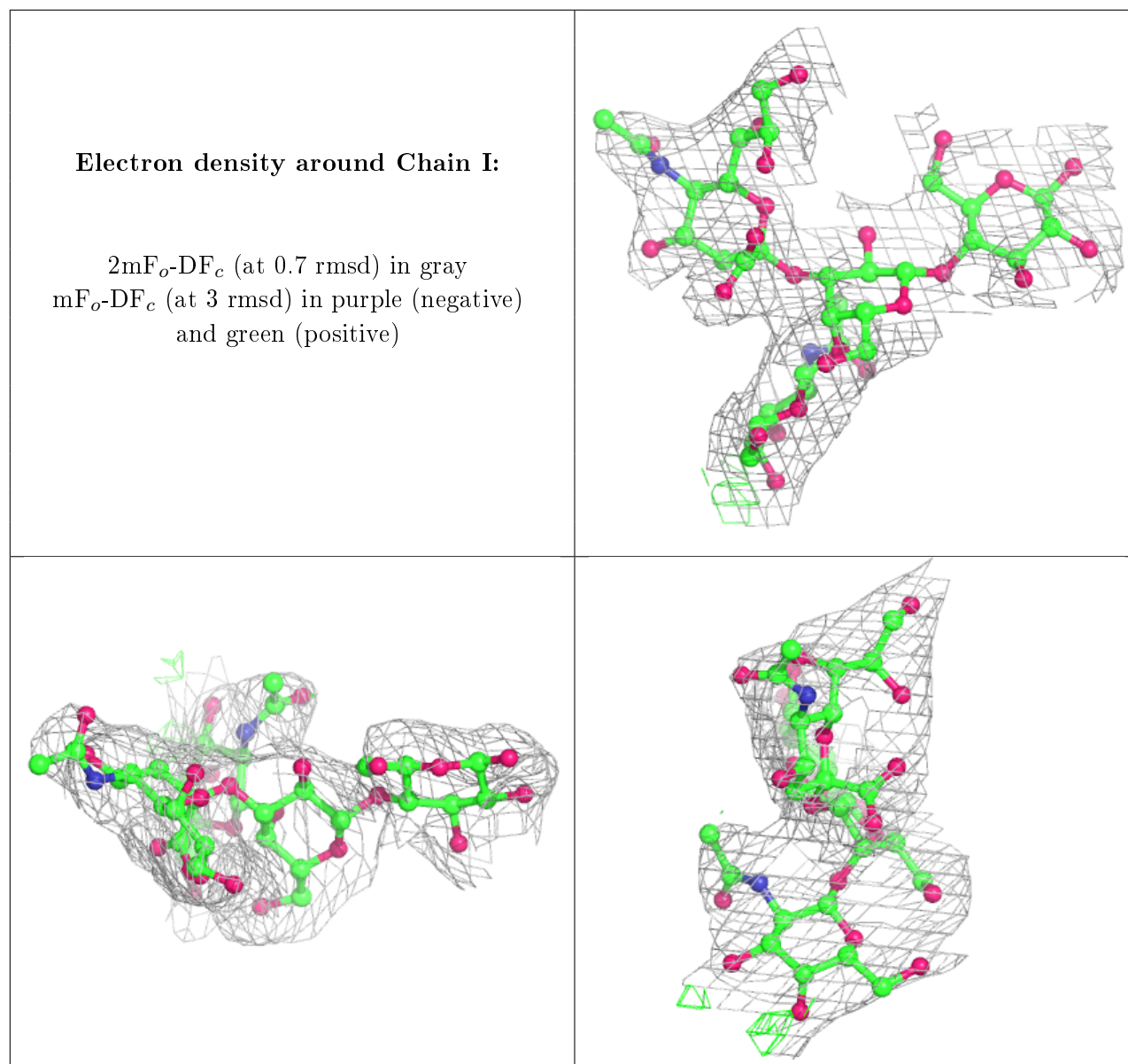


**Electron density around Chain H:**

$2mF_o-DF_c$  (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
6	GOL	D	1303	6/6	0.91	0.26	78,86,91,97	0
6	GOL	C	401	6/6	0.96	0.15	53,72,73,84	0
6	GOL	B	401	6/6	0.97	0.11	40,53,64,64	0
6	GOL	A	401	6/6	0.98	0.18	52,58,61,71	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.