



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2024 – 10:16 PM EST

PDB ID : 4MJ0  
Title : BK Polyomavirus VP1 pentamer in complex with GD3 oligosaccharide  
Authors : Neu, U.; Stroh, L.J.; Stehle, T.  
Deposited on : 2013-09-03  
Resolution : 1.70 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 : 2.36  
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) : 2.36

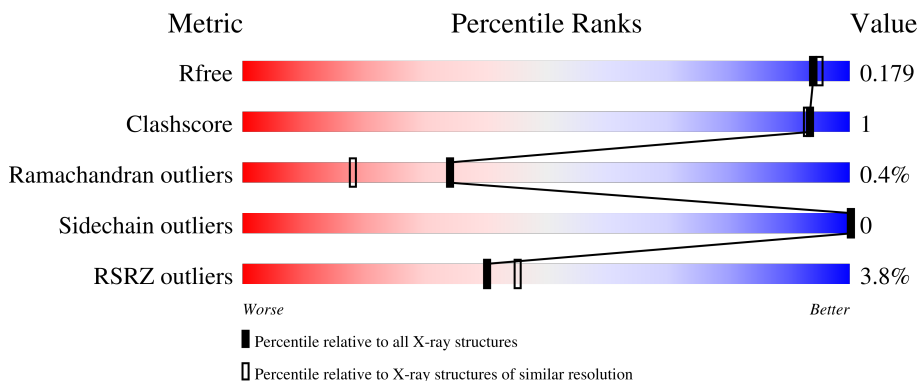
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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.





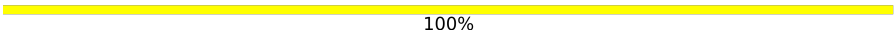
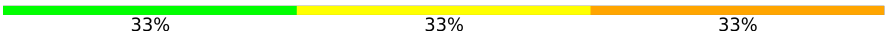
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	275	 94%
1	B	275	 93%
1	C	275	 85% 10%
1	D	275	 92%
1	E	275	 92%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	4	 25% 75%
3	G	2	 50% 50%
3	I	2	 100%
4	H	3	 33% 33% 33%

## 2 Entry composition i

There are 7 unique types of molecules in this entry. The entry contains 11705 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called VP1 capsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	265	Total 2062	C 1300	N 349	O 399	S 14	0	8	0
1	B	263	Total 2035	C 1279	N 348	O 394	S 14	0	7	0
1	C	247	Total 1916	C 1208	N 329	O 366	S 13	0	5	0
1	D	267	Total 2073	C 1305	N 356	O 399	S 13	0	6	0
1	E	264	Total 2095	C 1320	N 360	O 401	S 14	0	12	0

There are 20 discrepancies between the modelled and reference sequences:

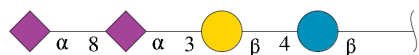
Chain	Residue	Modelled	Actual	Comment	Reference
A	26	GLY	-	expression tag	UNP Q85235
A	27	SER	-	expression tag	UNP Q85235
A	28	HIS	-	expression tag	UNP Q85235
A	29	MET	-	expression tag	UNP Q85235
B	26	GLY	-	expression tag	UNP Q85235
B	27	SER	-	expression tag	UNP Q85235
B	28	HIS	-	expression tag	UNP Q85235
B	29	MET	-	expression tag	UNP Q85235
C	26	GLY	-	expression tag	UNP Q85235
C	27	SER	-	expression tag	UNP Q85235
C	28	HIS	-	expression tag	UNP Q85235
C	29	MET	-	expression tag	UNP Q85235
D	26	GLY	-	expression tag	UNP Q85235
D	27	SER	-	expression tag	UNP Q85235
D	28	HIS	-	expression tag	UNP Q85235
D	29	MET	-	expression tag	UNP Q85235
E	26	GLY	-	expression tag	UNP Q85235
E	27	SER	-	expression tag	UNP Q85235
E	28	HIS	-	expression tag	UNP Q85235

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	29	MET	-	expression tag	UNP Q85235

- Molecule 2 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-8)-N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.



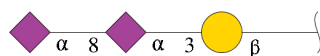
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	F	4	63	34	2	27	0	0	0

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-8)-N-acetyl-alpha-neuraminic acid.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	G	2	41	22	2	17	0	0	0
3	I	2	41	22	2	17	0	0	0

- Molecule 4 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-8)-N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	H	3	52	28	2	22	0	0	0

- Molecule 5 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
5	A	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	E	1	Total C O 6 3 3	0	0
5	E	1	Total C O 6 3 3	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Cl 1 1	0	0
6	B	1	Total Cl 1 1	0	0
6	C	1	Total Cl 1 1	0	0
6	D	1	Total Cl 1 1	0	0
6	E	1	Total Cl 1 1	0	0

- Molecule 7 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
7	A	300	Total 300	O 300	0	0
7	B	245	Total 245	O 245	0	0
7	C	192	Total 192	O 192	0	0
7	D	266	Total 266	O 266	0	0
7	E	277	Total 277	O 277	0	0





- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-8)-N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain F:  25% 75%

  
BGCL1  
GAL2  
SIA3  
SIA4

- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-8)-N-acetyl-alpha-neuraminic acid

Chain G:  50% 50%

  
SIA1  
SIA2

- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-8)-N-acetyl-alpha-neuraminic acid

Chain I:  100%

  
SIA1  
SIA2

- Molecule 4: N-acetyl-alpha-neuraminic acid-(2-8)-N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose

Chain H:  33% 33% 33%

  
GAL1  
SIA2  
SIA3

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	144.68Å 152.64Å 63.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.00 – 1.70 48.00 – 1.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.00-1.70) 100.0 (48.00-1.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.72 (at 1.70Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.150 , 0.177 0.156 , 0.179	Depositor DCC
$R_{free}$ test set	5395 reflections (3.50%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.2	Xtrriage
Anisotropy	0.260	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 40.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	11705	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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: SIA, GOL, GAL, CL, BGC

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	0.48	0/2132	0.66	0/2902
1	B	0.45	0/2101	0.62	0/2862
1	C	0.44	0/1975	0.61	0/2686
1	D	0.49	0/2138	0.66	0/2911
1	E	0.51	0/2177	0.69	0/2957
All	All	0.47	0/10523	0.65	0/14318

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [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	A	2062	0	2013	5	0
1	B	2035	0	1963	4	0
1	C	1916	0	1852	9	0
1	D	2073	0	2029	7	0
1	E	2095	0	2073	7	0
2	F	63	0	53	0	0
3	G	41	0	34	0	0
3	I	41	0	34	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	H	52	0	44	1	0
5	A	6	0	8	0	0
5	B	12	0	16	0	0
5	C	6	0	8	0	0
5	D	6	0	8	0	0
5	E	12	0	16	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	E	1	0	0	0	0
7	A	300	0	0	1	0
7	B	245	0	0	0	0
7	C	192	0	0	0	0
7	D	266	0	0	0	0
7	E	277	0	0	1	0
All	All	11705	0	10151	29	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 1.

All (29) 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:C:109:MET:CE	1:C:295:VAL:HG11	2.11	0.80
1:C:109:MET:HE2	1:C:295:VAL:HG11	1.67	0.74
1:C:113:VAL:HG12	1:C:114:THR:HG23	1.85	0.59
1:E:35:LEU:HD11	1:E:296:LYS:HG2	1.86	0.56
1:C:259:LEU:HD21	1:C:288:ILE:HD13	1.88	0.56
1:A:259:LEU:HD21	1:A:288:ILE:HD13	1.89	0.53
1:A:68:LYS:HD3	7:A:775:HOH:O	2.10	0.51
1:C:255:LYS:O	1:C:256:ALA:HB3	2.14	0.48
1:E:98:LEU:HD21	1:E:110:TRP:CZ2	2.48	0.48
1:D:48:GLU:OE1	1:D:285:TYR:OH	2.28	0.48
1:E:259:LEU:HD21	1:E:288:ILE:HD13	1.96	0.47
1:D:188:VAL:HG23	1:D:189:MET:H	1.80	0.46
1:B:106:ASN:O	1:B:107:LEU:HD23	2.16	0.45
1:C:109:MET:HE2	1:C:295:VAL:CG1	2.41	0.45
1:C:231:PRO:HB3	1:D:226:GLY:O	2.18	0.43
1:D:98:LEU:HD21	1:D:110:TRP:CZ2	2.54	0.43
1:D:46:GLU:HG2	1:D:289:ARG:HG2	2.01	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:58:PRO:HA	1:B:83:LYS:HD3	2.01	0.42
1:A:231:PRO:HB3	1:B:230[A]:VAL:HG11	2.00	0.42
1:E:103:THR:O	1:E:103:THR:HG22	2.20	0.42
1:D:43:ALA:HB3	1:D:292:LYS:HE3	2.01	0.42
1:D:142:LYS:HE3	1:D:145[B]:GLN:OE1	2.20	0.42
1:C:109:MET:HE3	1:C:253:LEU:HD13	2.02	0.42
1:A:231:PRO:HB3	1:B:226:GLY:O	2.20	0.41
1:E:188:VAL:HG23	1:E:189:MET:H	1.86	0.41
1:A:98:LEU:HD21	1:A:292:LYS:HE2	2.02	0.41
1:C:276:THR:CG2	4:H:2:SIA:H112	2.51	0.41
1:E:172:TYR:CG	1:E:178:THR:HG21	2.55	0.41
1:E:116[B]:GLN:HG3	7:E:680:HOH:O	2.21	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	271/275 (98%)	258 (95%)	12 (4%)	1 (0%)	34	18
1	B	266/275 (97%)	255 (96%)	10 (4%)	1 (0%)	34	18
1	C	246/275 (90%)	236 (96%)	9 (4%)	1 (0%)	34	18
1	D	271/275 (98%)	260 (96%)	10 (4%)	1 (0%)	34	18
1	E	272/275 (99%)	261 (96%)	10 (4%)	1 (0%)	34	18
All	All	1326/1375 (96%)	1270 (96%)	51 (4%)	5 (0%)	34	18

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	188	VAL
1	C	188	VAL

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	188	VAL
1	B	188	VAL
1	D	188	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	230/236 (98%)	230 (100%)	0	100	100
1	B	225/236 (95%)	225 (100%)	0	100	100
1	C	211/236 (89%)	211 (100%)	0	100	100
1	D	232/236 (98%)	232 (100%)	0	100	100
1	E	238/236 (101%)	238 (100%)	0	100	100
All	All	1136/1180 (96%)	1136 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

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
2	BGC	F	1	2	12,12,12	0.46	0	17,17,17	0.65	0
2	GAL	F	2	2	11,11,12	0.48	0	15,15,17	1.11	1 (6%)
2	SIA	F	3	2	20,20,21	0.61	0	24,28,31	1.34	4 (16%)
2	SIA	F	4	2	20,20,21	0.62	0	24,28,31	1.21	3 (12%)
3	SIA	G	1	3	21,21,21	1.00	1 (4%)	25,31,31	1.31	4 (16%)
3	SIA	G	2	3	20,20,21	0.60	0	24,28,31	0.84	0
4	GAL	H	1	4	12,12,12	0.50	0	17,17,17	0.74	0
4	SIA	H	2	4	20,20,21	0.62	0	24,28,31	1.51	5 (20%)
4	SIA	H	3	4	20,20,21	0.59	0	24,28,31	1.17	3 (12%)
3	SIA	I	1	3	21,21,21	1.03	1 (4%)	25,31,31	1.37	3 (12%)
3	SIA	I	2	3	20,20,21	0.67	0	24,28,31	1.03	3 (12%)

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	BGC	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	SIA	F	3	2	-	2/18/34/38	0/1/1/1
2	SIA	F	4	2	-	1/18/34/38	0/1/1/1
3	SIA	G	1	3	-	5/20/38/38	0/1/1/1
3	SIA	G	2	3	-	0/18/34/38	0/1/1/1
4	GAL	H	1	4	-	0/2/22/22	0/1/1/1
4	SIA	H	2	4	-	2/18/34/38	0/1/1/1
4	SIA	H	3	4	-	0/18/34/38	0/1/1/1
3	SIA	I	1	3	-	6/20/38/38	0/1/1/1
3	SIA	I	2	3	-	0/18/34/38	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	1	SIA	O2-C2	3.53	1.44	1.39

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	1	SIA	O2-C2	3.27	1.43	1.39

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	1	SIA	O1A-C1-C2	-3.73	117.94	123.59
3	G	1	SIA	O1A-C1-C2	-3.54	118.23	123.59
4	H	2	SIA	C6-O6-C2	3.27	118.34	111.34
2	F	3	SIA	O6-C2-C1	3.22	114.01	107.70
3	G	1	SIA	O6-C6-C5	2.96	112.67	109.78
2	F	3	SIA	O1B-C1-C2	2.94	121.41	113.03
4	H	2	SIA	O6-C2-C3	-2.86	106.52	110.46
2	F	2	GAL	O5-C5-C6	2.85	111.68	107.20
2	F	4	SIA	O6-C2-C3	-2.75	106.67	110.46
2	F	4	SIA	C6-O6-C2	2.58	116.85	111.34
2	F	3	SIA	C6-O6-C2	2.54	116.76	111.34
3	I	1	SIA	O8-C8-C9	-2.53	103.20	109.14
4	H	2	SIA	O6-C2-C1	2.42	112.44	107.70
3	G	1	SIA	O8-C8-C7	-2.33	103.44	109.10
3	I	2	SIA	O1B-C1-C2	2.29	119.56	113.03
4	H	2	SIA	O1B-C1-C2	2.25	119.44	113.03
4	H	3	SIA	C6-O6-C2	2.25	116.14	111.34
3	I	2	SIA	O1A-C1-C2	-2.22	117.33	122.57
4	H	2	SIA	C3-C4-C5	2.22	114.14	111.46
3	I	2	SIA	C6-O6-C2	2.20	116.04	111.34
2	F	4	SIA	C8-C7-C6	-2.18	108.91	113.03
2	F	3	SIA	O1A-C1-C2	-2.14	117.52	122.57
4	H	3	SIA	C8-C7-C6	-2.06	109.13	113.03
3	I	1	SIA	O2-C2-C1	-2.05	106.58	110.76
4	H	3	SIA	O1B-C1-C2	2.01	118.78	113.03
3	G	1	SIA	C3-C4-C5	2.01	113.06	109.98

There are no chirality outliers.

All (16) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	1	SIA	O6-C6-C7-O7
3	I	1	SIA	C6-C7-C8-C9
3	I	1	SIA	C6-C7-C8-O8
3	I	1	SIA	O7-C7-C8-C9
3	I	1	SIA	O7-C7-C8-O8
4	H	2	SIA	C7-C8-C9-O9

*Continued on next page...*



*Continued from previous page...*

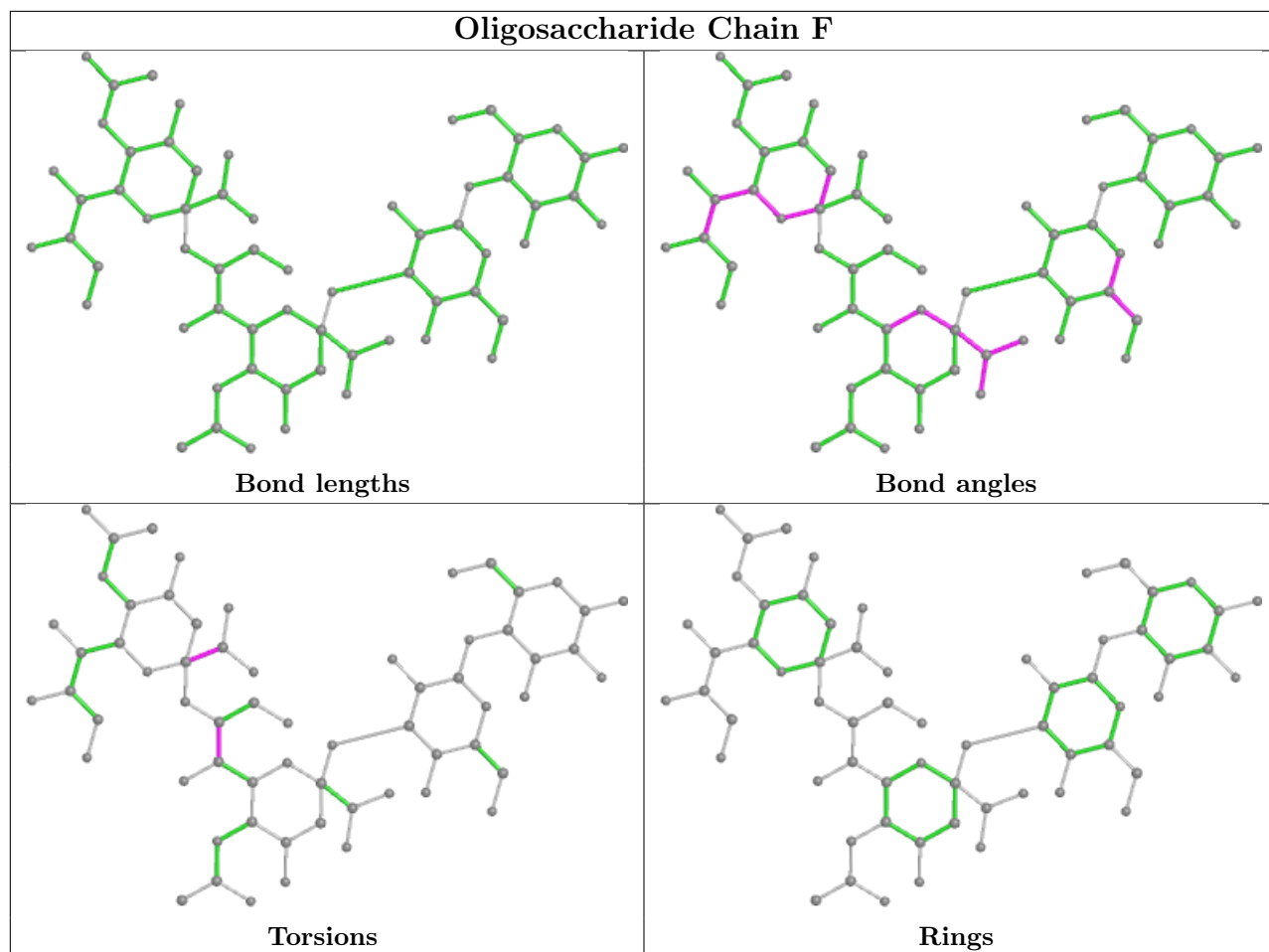
Mol	Chain	Res	Type	Atoms
3	G	1	SIA	C7-C8-C9-O9
4	H	2	SIA	O8-C8-C9-O9
3	G	1	SIA	O8-C8-C9-O9
2	F	3	SIA	C6-C7-C8-C9
2	F	3	SIA	C6-C7-C8-O8
3	I	1	SIA	O1B-C1-C2-C3
2	F	4	SIA	O1A-C1-C2-O6
3	G	1	SIA	O1A-C1-C2-O2
3	I	1	SIA	O1A-C1-C2-O2
3	G	1	SIA	O1B-C1-C2-O6

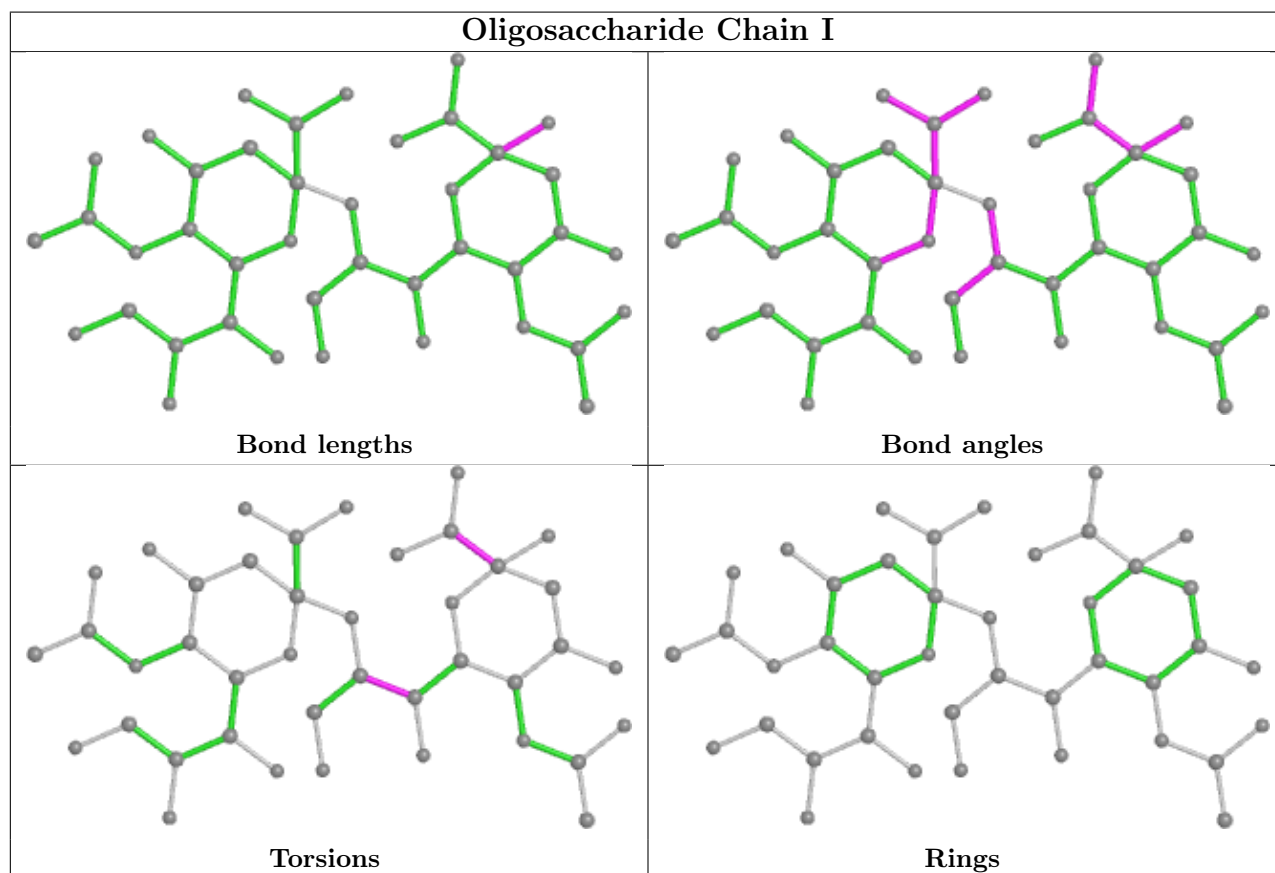
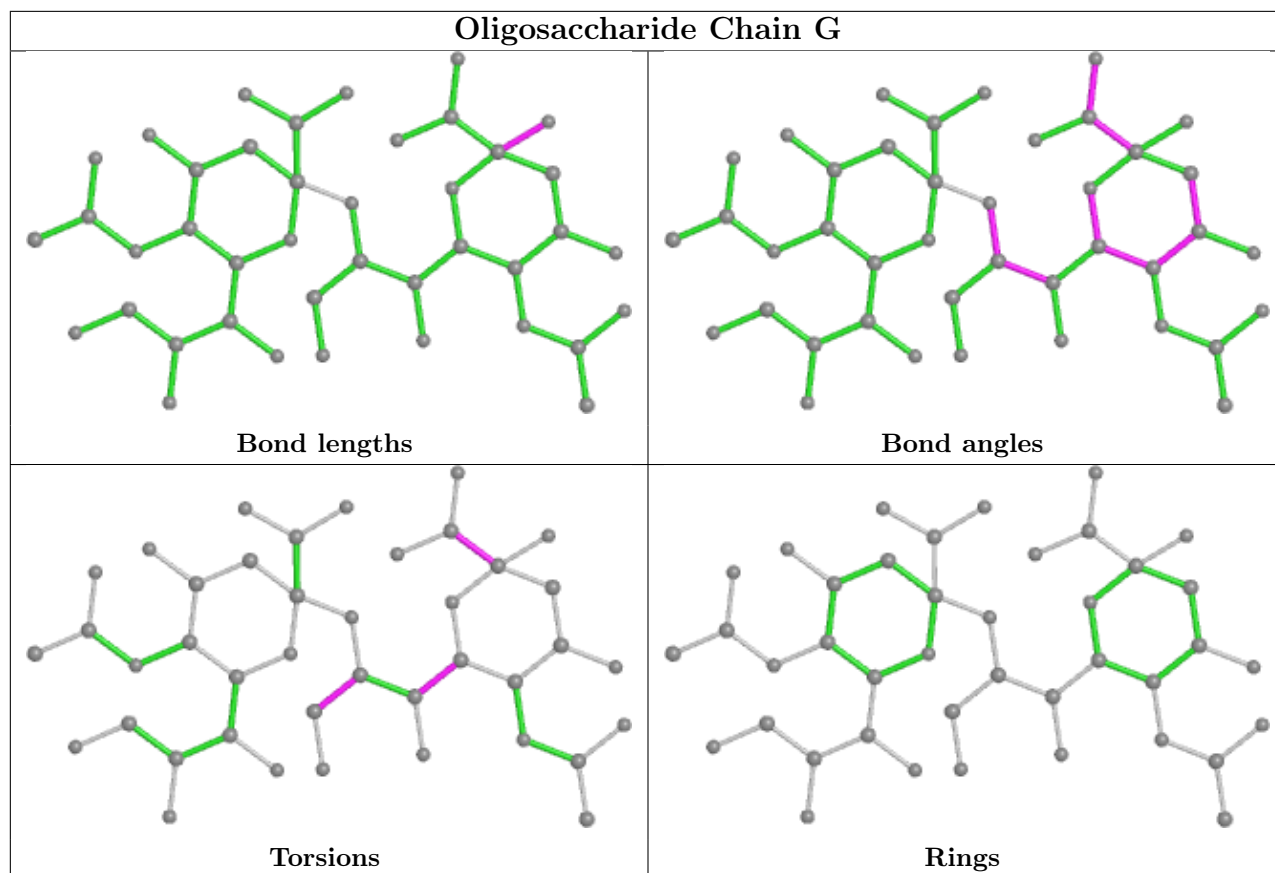
There are no ring outliers.

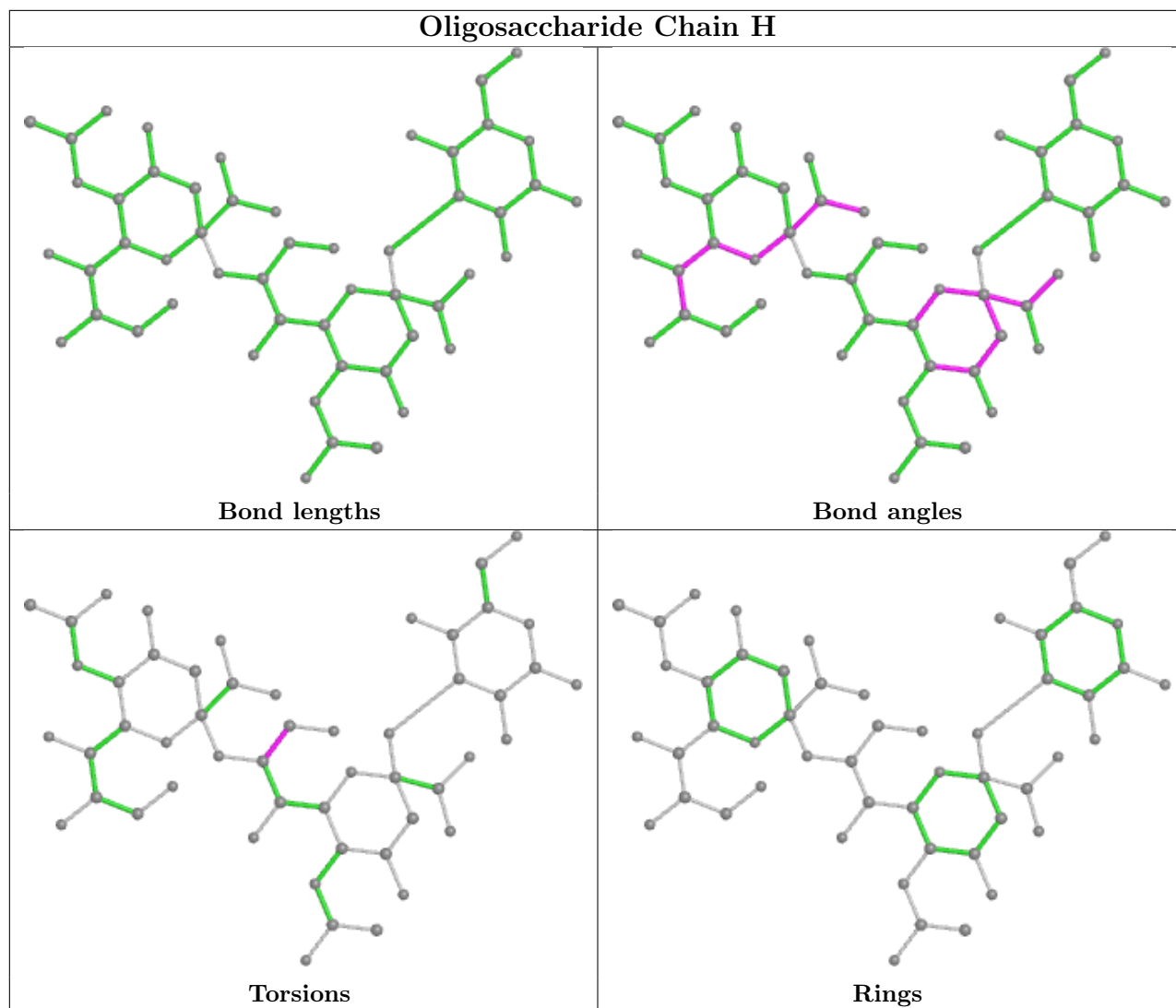
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	2	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](#)

Of 12 ligands modelled in this entry, 5 are monoatomic - leaving 7 for Mogul analysis.

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	GOL	A	405	-	5,5,5	0.39	0	5,5,5	0.11	0
5	GOL	B	404	-	5,5,5	0.39	0	5,5,5	0.30	0
5	GOL	B	403	-	5,5,5	0.44	0	5,5,5	0.25	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	GOL	E	402	-	5,5,5	0.39	0	5,5,5	0.38	0
5	GOL	E	401	-	5,5,5	0.42	0	5,5,5	0.23	0
5	GOL	D	403	-	5,5,5	0.42	0	5,5,5	0.26	0
5	GOL	C	404	-	5,5,5	0.43	0	5,5,5	0.36	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	GOL	A	405	-	-	4/4/4/4	-
5	GOL	B	404	-	-	4/4/4/4	-
5	GOL	B	403	-	-	2/4/4/4	-
5	GOL	E	402	-	-	2/4/4/4	-
5	GOL	E	401	-	-	4/4/4/4	-
5	GOL	D	403	-	-	2/4/4/4	-
5	GOL	C	404	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (22) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	405	GOL	O1-C1-C2-C3
5	A	405	GOL	C1-C2-C3-O3
5	B	404	GOL	O1-C1-C2-C3
5	B	404	GOL	C1-C2-C3-O3
5	C	404	GOL	O1-C1-C2-C3
5	E	401	GOL	O1-C1-C2-C3
5	C	404	GOL	O1-C1-C2-O2
5	B	403	GOL	C1-C2-C3-O3
5	C	404	GOL	C1-C2-C3-O3
5	D	403	GOL	C1-C2-C3-O3
5	E	402	GOL	C1-C2-C3-O3
5	B	403	GOL	O2-C2-C3-O3
5	B	404	GOL	O1-C1-C2-O2
5	B	404	GOL	O2-C2-C3-O3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
5	D	403	GOL	O2-C2-C3-O3
5	A	405	GOL	O1-C1-C2-O2
5	E	402	GOL	O2-C2-C3-O3
5	A	405	GOL	O2-C2-C3-O3
5	C	404	GOL	O2-C2-C3-O3
5	E	401	GOL	O1-C1-C2-O2
5	E	401	GOL	O2-C2-C3-O3
5	E	401	GOL	C1-C2-C3-O3

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [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	265/275 (96%)	-0.28	4 (1%) 73 77	11, 17, 33, 48	0
1	B	263/275 (95%)	-0.08	15 (5%) 23 26	12, 20, 56, 75	0
1	C	247/275 (89%)	-0.04	16 (6%) 18 21	13, 22, 60, 87	0
1	D	267/275 (97%)	-0.28	1 (0%) 92 93	10, 17, 37, 60	0
1	E	264/275 (96%)	-0.19	13 (4%) 29 33	10, 16, 44, 90	0
All	All	1306/1375 (94%)	-0.17	49 (3%) 40 45	10, 18, 45, 90	0

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	104	CYS	9.9
1	B	102	LEU	9.0
1	B	104	CYS	7.7
1	E	103	THR	7.0
1	B	103	THR	6.2
1	C	107	LEU	5.9
1	C	295	VAL	5.8
1	C	37	VAL	5.5
1	B	105	GLY	4.9
1	C	34	VAL	4.9
1	E	298	PRO	4.8
1	C	108	LEU	4.5
1	E	44	ILE	4.4
1	E	105	GLY	4.2
1	C	255	LYS	4.2
1	B	106	ASN	3.7
1	B	101	ASP	3.7
1	C	44	ILE	3.6
1	B	39	THR	3.5
1	E	102	LEU	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	100	GLU	3.2
1	B	255	LYS	3.2
1	C	35	LEU	3.1
1	E	101	ASP	3.0
1	C	110	TRP	3.0
1	B	107	LEU	2.8
1	C	256	ALA	2.7
1	C	294	SER	2.7
1	C	97	ASN	2.7
1	E	297	ASN	2.7
1	C	248	GLN	2.7
1	B	297	ASN	2.6
1	C	109	MET	2.5
1	A	32	VAL	2.5
1	A	107	LEU	2.5
1	E	106	ASN	2.5
1	B	44	ILE	2.4
1	B	36	GLU	2.3
1	C	292	LYS	2.3
1	E	32	VAL	2.2
1	B	34	VAL	2.1
1	C	247	GLU	2.1
1	E	107	LEU	2.1
1	B	37	VAL	2.1
1	A	102	LEU	2.1
1	A	103	THR	2.1
1	D	37	VAL	2.1
1	E	296	LYS	2.1
1	B	35	LEU	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.

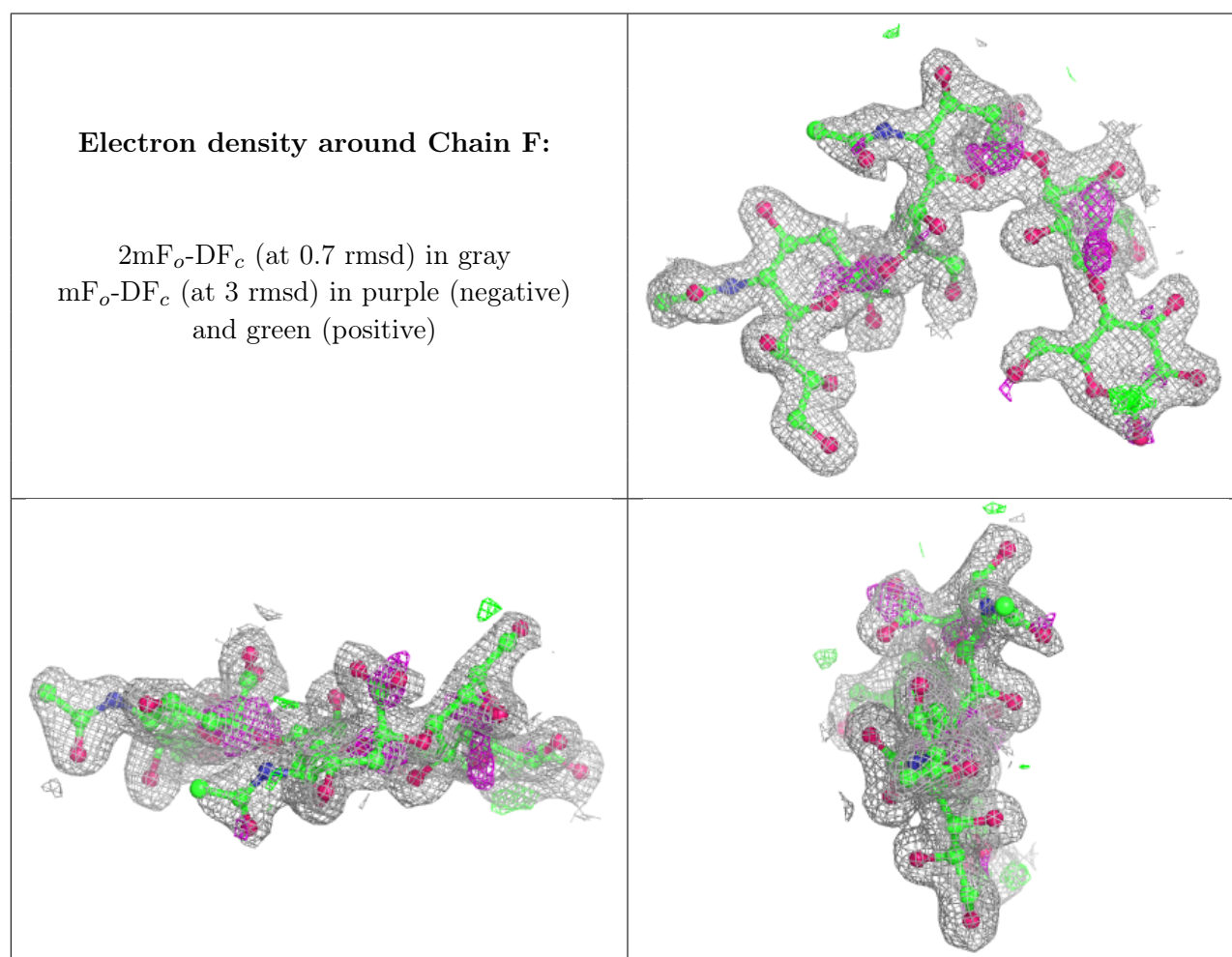
## 6.3 Carbohydrates [i](#)

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.



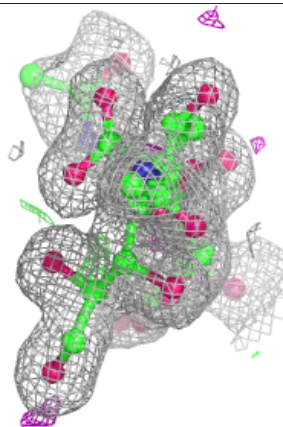
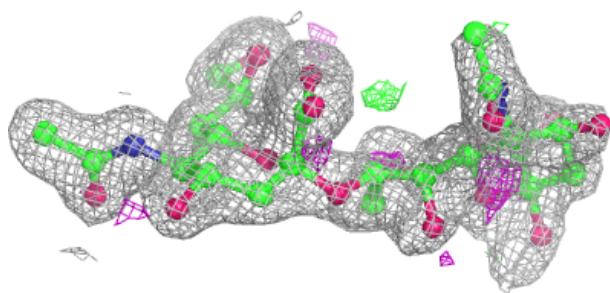
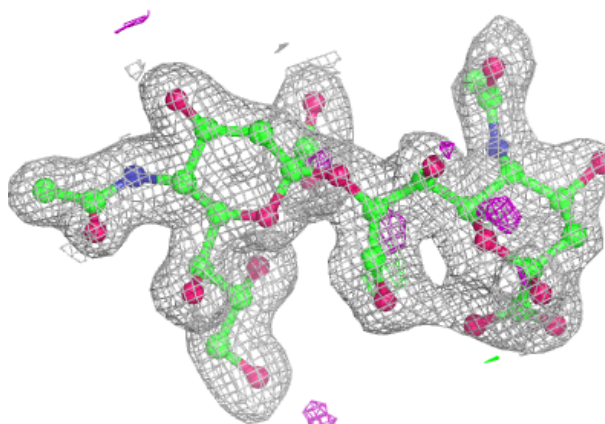
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SIA	H	2	20/21	0.74	0.30	35,42,49,50	0
4	GAL	H	1	12/12	0.77	0.18	38,40,41,43	12
3	SIA	G	1	21/21	0.84	0.29	31,45,51,54	0
2	SIA	F	3	20/21	0.86	0.28	29,32,39,41	0
2	BGC	F	1	12/12	0.88	0.28	30,38,43,44	0
3	SIA	I	1	21/21	0.88	0.17	21,31,34,36	0
2	SIA	F	4	20/21	0.90	0.17	21,27,32,32	0
4	SIA	H	3	20/21	0.91	0.20	25,29,33,35	0
2	GAL	F	2	11/12	0.92	0.26	25,30,35,41	0
3	SIA	G	2	20/21	0.93	0.11	20,25,29,34	0
3	SIA	I	2	20/21	0.94	0.10	17,19,22,22	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.



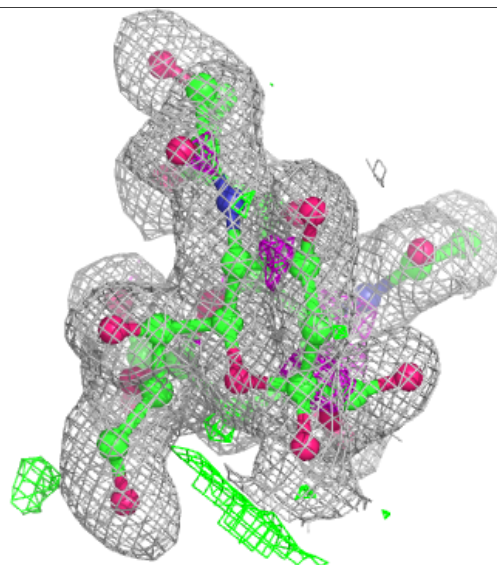
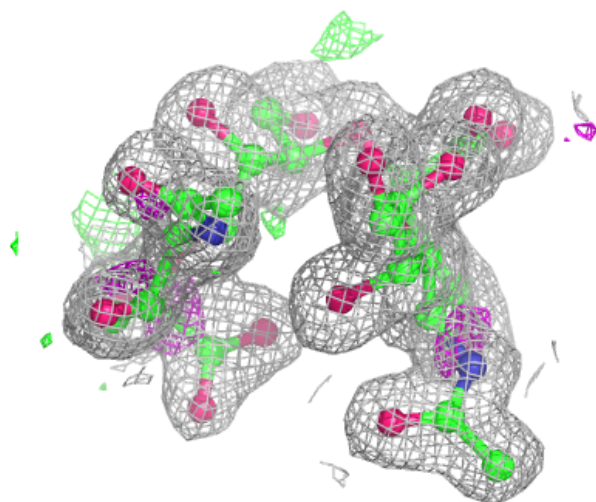
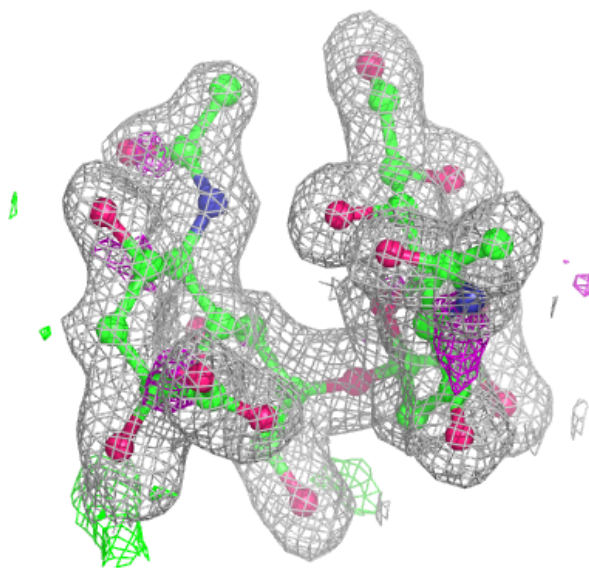
**Electron density around Chain G:**

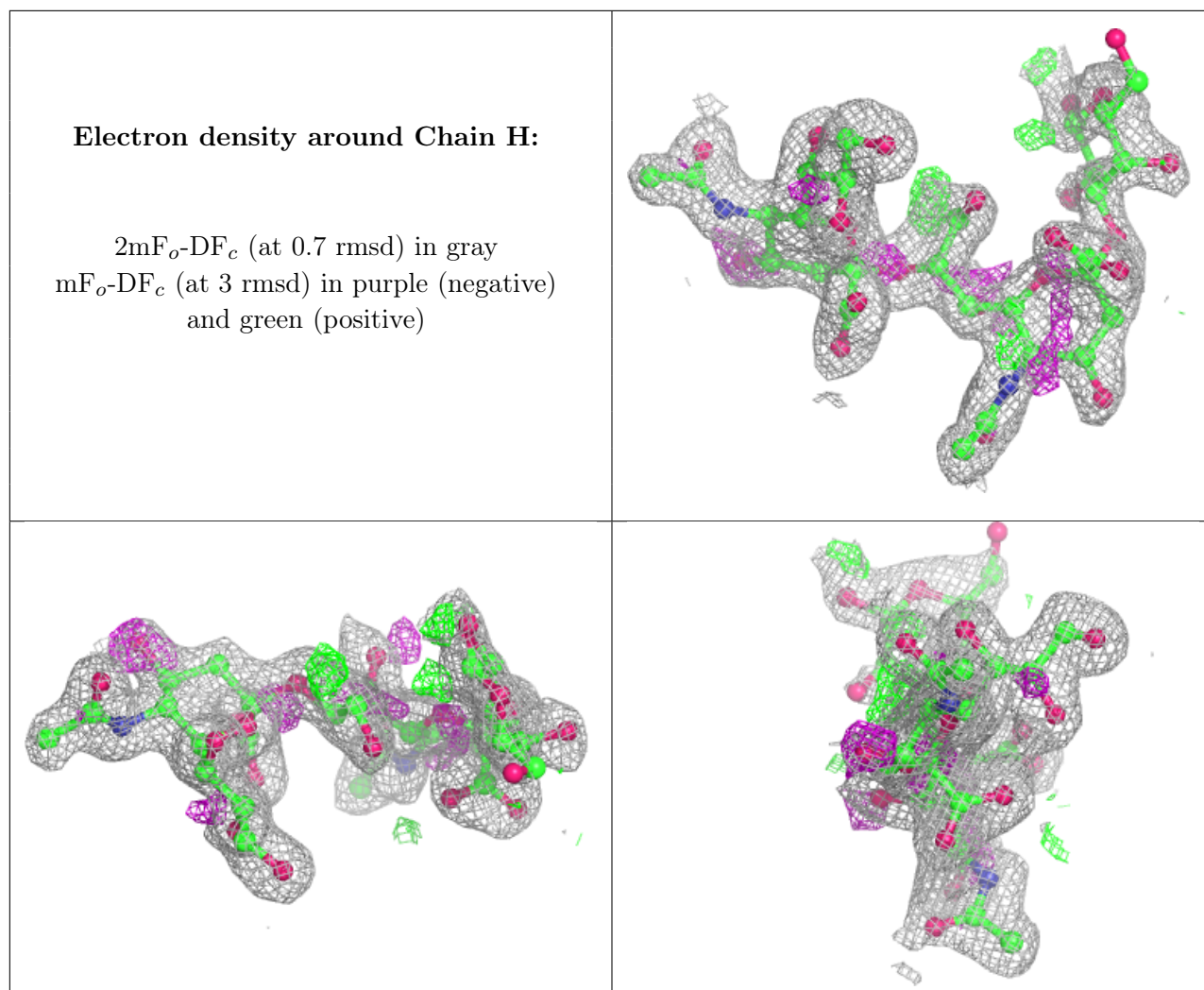
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain I:**

$2mF_o-DF_c$  (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	GOL	D	403	6/6	0.80	0.14	36,44,46,48	0
5	GOL	B	404	6/6	0.85	0.13	32,41,45,48	0
5	GOL	A	405	6/6	0.85	0.12	33,39,42,42	0
5	GOL	E	401	6/6	0.86	0.23	37,42,44,45	0
5	GOL	B	403	6/6	0.87	0.29	40,43,44,46	0
5	GOL	C	404	6/6	0.90	0.10	32,41,43,46	0
5	GOL	E	402	6/6	0.90	0.10	30,41,43,47	0
6	CL	A	406	1/1	0.95	0.06	26,26,26,26	0
6	CL	C	405	1/1	0.96	0.06	35,35,35,35	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	CL	D	404	1/1	0.97	0.05	27,27,27,27	0
6	CL	B	405	1/1	0.98	0.05	28,28,28,28	0
6	CL	E	403	1/1	0.99	0.05	28,28,28,28	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.