



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 12, 2023 – 08:15 AM EDT

PDB ID : 4N64  
Title : Crystal structure of hemagglutinin from an H7N9 influenza virus in complex with a biantennary glycan receptor  
Authors : Xu, R.; Wilson, I.A.  
Deposited on : 2013-10-11  
Resolution : 2.70 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

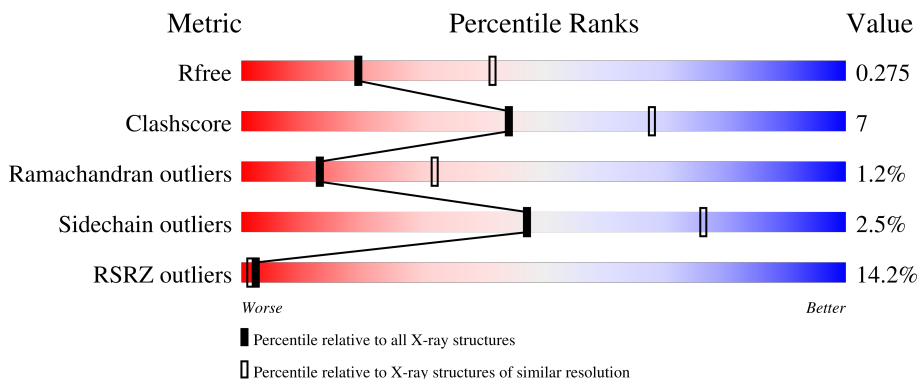
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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.



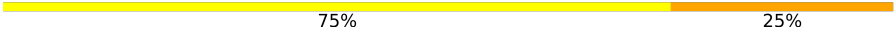

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	321	
1	C	321	
2	B	183	
2	D	183	
3	E	2	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	F	4	 75% 25%
5	G	2	 50% 50%

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 7885 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 Hemagglutinin HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	318	2427	1509	438	465	15	0	0	0
1	C	318	2427	1509	438	465	15	0	0	0

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	169	1373	845	241	280	7	0	0	0
2	D	169	1373	845	241	280	7	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	177	SER	-	expression tag	UNP R4NN21
B	178	GLY	-	expression tag	UNP R4NN21
B	179	ARG	-	expression tag	UNP R4NN21
B	180	LEU	-	expression tag	UNP R4NN21
B	181	VAL	-	expression tag	UNP R4NN21
B	182	PRO	-	expression tag	UNP R4NN21
B	183	ARG	-	expression tag	UNP R4NN21
D	177	SER	-	expression tag	UNP R4NN21
D	178	GLY	-	expression tag	UNP R4NN21
D	179	ARG	-	expression tag	UNP R4NN21
D	180	LEU	-	expression tag	UNP R4NN21
D	181	VAL	-	expression tag	UNP R4NN21
D	182	PRO	-	expression tag	UNP R4NN21
D	183	ARG	-	expression tag	UNP R4NN21

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-a

cetamido-2-deoxy-beta-D-glucopyranose.



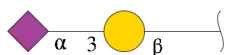
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	E	2	28	16	2	10	0	0	0

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	F	4	60	33	3	24	0	0	0

- Molecule 5 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	G	2	31	17	1	13	0	0	0

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

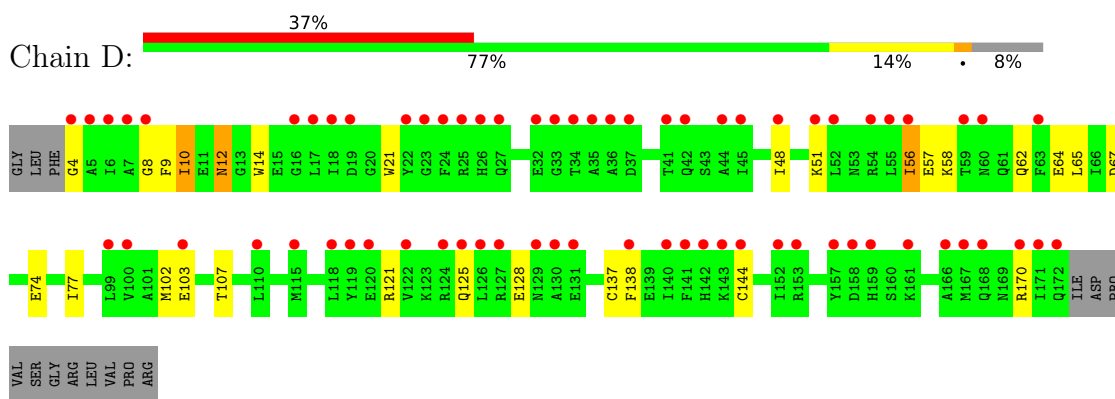


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	A	1	14	8	1	5	0	0
6	B	1	14	8	1	5	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	83	83	83	0	0
7	B	43	43	43	0	0
7	C	9	9	9	0	0
7	D	3	3	3	0	0

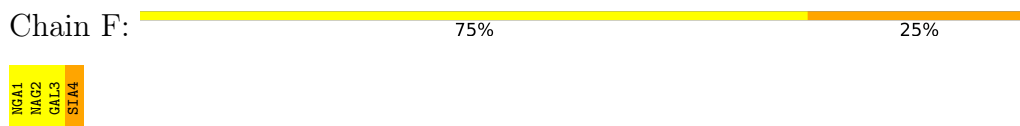




- Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



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



- Molecule 5: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	153.86Å 153.86Å 153.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.66 – 2.70 48.66 – 2.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.66-2.70) 99.8 (48.66-2.70)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 2.69Å)	Xtrriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, $R_{free}$	0.226 , 0.274 0.227 , 0.275	Depositor DCC
$R_{free}$ test set	1695 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	54.4	Xtrriage
Anisotropy	0.000	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 50.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.032 for l,-k,h	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7885	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.94% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, NGA, GAL, SIA

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	0.24	0/2474	0.42	0/3345
1	C	0.23	0/2474	0.40	0/3345
2	B	0.24	0/1396	0.40	0/1881
2	D	0.22	0/1396	0.38	0/1881
All	All	0.23	0/7740	0.40	0/10452

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	2427	0	2387	36	0
1	C	2427	0	2389	45	0
2	B	1373	0	1273	13	0
2	D	1373	0	1274	25	0
3	E	28	0	25	1	0
4	F	60	0	52	2	0
5	G	31	0	26	0	0
6	A	14	0	13	0	0
6	B	14	0	13	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	83	0	0	2	0
7	B	43	0	0	0	0
7	C	9	0	0	0	0
7	D	3	0	0	1	0
All	All	7885	0	7452	100	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

The worst 5 of 100 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:291:ASN:OD1	2:D:58:LYS:NZ	1.97	0.97
1:C:195:TYR:O	1:C:200:LYS:NZ	2.08	0.85
1:A:56:LYS:NZ	1:A:279:GLY:O	2.09	0.84
1:C:56:LYS:NZ	1:C:279:GLY:O	2.16	0.79
1:A:158(A):ASP:OD2	1:A:193:LYS:NZ	2.17	0.77

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	316/321 (98%)	295 (93%)	18 (6%)	3 (1%)	17	40
1	C	316/321 (98%)	295 (93%)	18 (6%)	3 (1%)	17	40
2	B	167/183 (91%)	155 (93%)	9 (5%)	3 (2%)	8	21
2	D	167/183 (91%)	155 (93%)	9 (5%)	3 (2%)	8	21
All	All	966/1008 (96%)	900 (93%)	54 (6%)	12 (1%)	13	32

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	12	ASN
2	D	12	ASN
1	A	157	ASN
1	C	157	ASN
2	D	57	GLU

### 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	267/269 (99%)	258 (97%)	9 (3%)	37	66
1	C	267/269 (99%)	259 (97%)	8 (3%)	41	70
2	B	145/157 (92%)	143 (99%)	2 (1%)	67	86
2	D	145/157 (92%)	143 (99%)	2 (1%)	67	86
All	All	824/852 (97%)	803 (98%)	21 (2%)	47	76

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	78	GLN
1	C	211	GLN
2	D	64	GLU
1	C	237	LEU
1	C	158(A)	ASP

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

Mol	Chain	Res	Type
2	D	12	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](#)

8 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	E	1	3,1	14,14,15	0.54	0	17,19,21	0.45	0
3	NAG	E	2	3	14,14,15	0.29	0	17,19,21	0.46	0
4	NGA	F	1	4	15,15,15	0.44	0	21,21,21	1.65	4 (19%)
4	NAG	F	2	4	14,14,15	0.28	0	17,19,21	0.44	0
4	GAL	F	3	4	11,11,12	1.05	2 (18%)	15,15,17	1.27	1 (6%)
4	SIA	F	4	4	20,20,21	0.74	0	24,28,31	1.59	4 (16%)
5	GAL	G	1	5	11,11,12	0.85	0	15,15,17	0.89	0
5	SIA	G	2	5	20,20,21	0.70	0	24,28,31	1.55	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
3	NAG	E	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	2/6/23/26	0/1/1/1
4	NGA	F	1	4	-	4/6/26/26	0/1/1/1
4	NAG	F	2	4	-	0/6/23/26	0/1/1/1
4	GAL	F	3	4	-	0/2/19/22	0/1/1/1
4	SIA	F	4	4	-	4/18/34/38	0/1/1/1
5	GAL	G	1	5	-	0/2/19/22	0/1/1/1
5	SIA	G	2	5	-	6/18/34/38	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	3	GAL	C2-C3	2.19	1.55	1.52
4	F	3	GAL	C1-C2	2.04	1.56	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	4	SIA	C6-O6-C2	5.27	122.62	111.34
5	G	2	SIA	C6-O6-C2	4.98	122.00	111.34
4	F	1	NGA	O5-C1-C2	4.33	113.87	109.52
4	F	1	NGA	C1-C2-N2	-3.88	106.23	110.73
4	F	3	GAL	C1-C2-C3	3.56	114.05	109.67

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

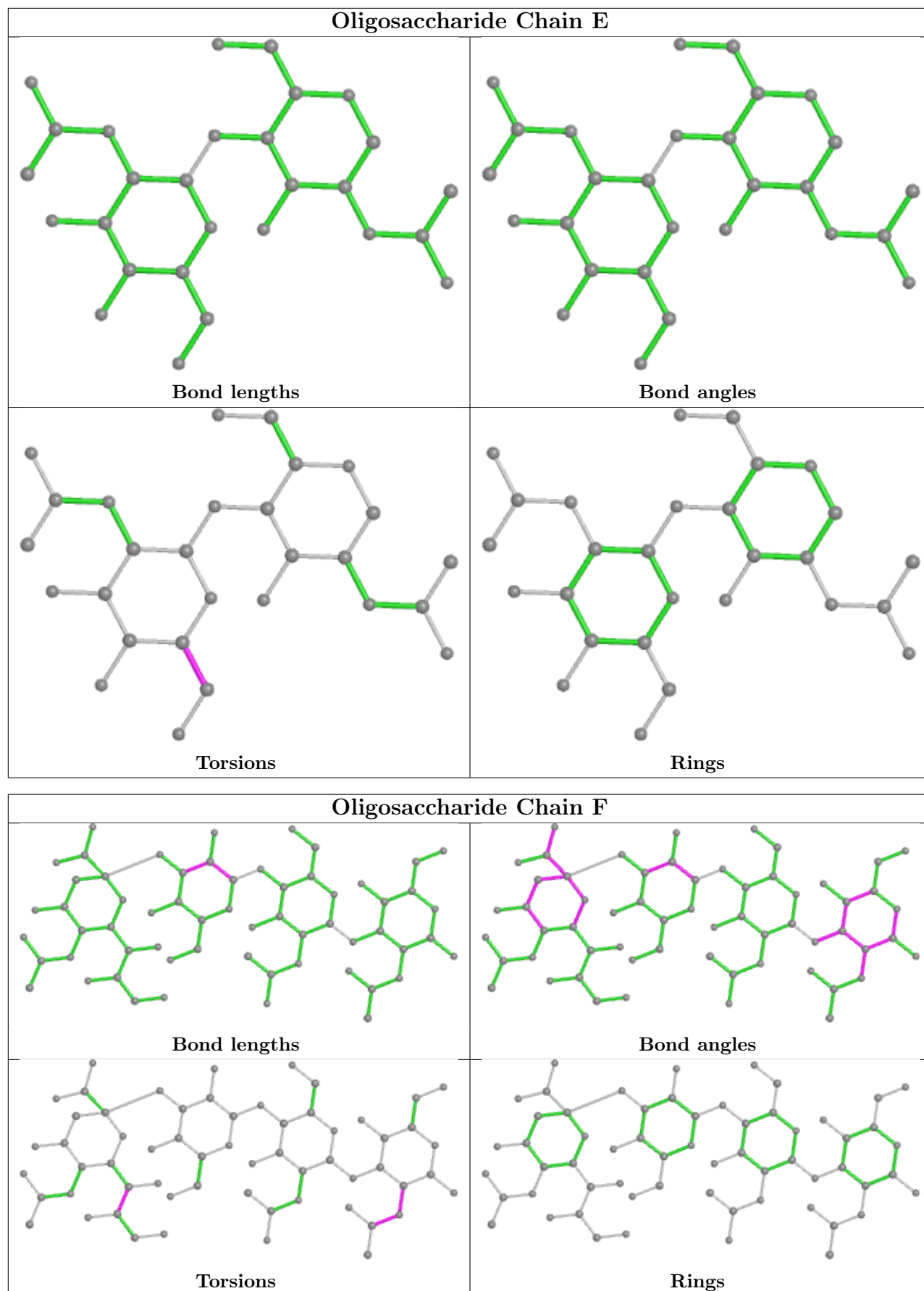
Mol	Chain	Res	Type	Atoms
4	F	1	NGA	C1-C2-N2-C7
4	F	1	NGA	C3-C2-N2-C7
4	F	1	NGA	C8-C7-N2-C2
4	F	1	NGA	O7-C7-N2-C2
4	F	4	SIA	C6-C7-C8-C9

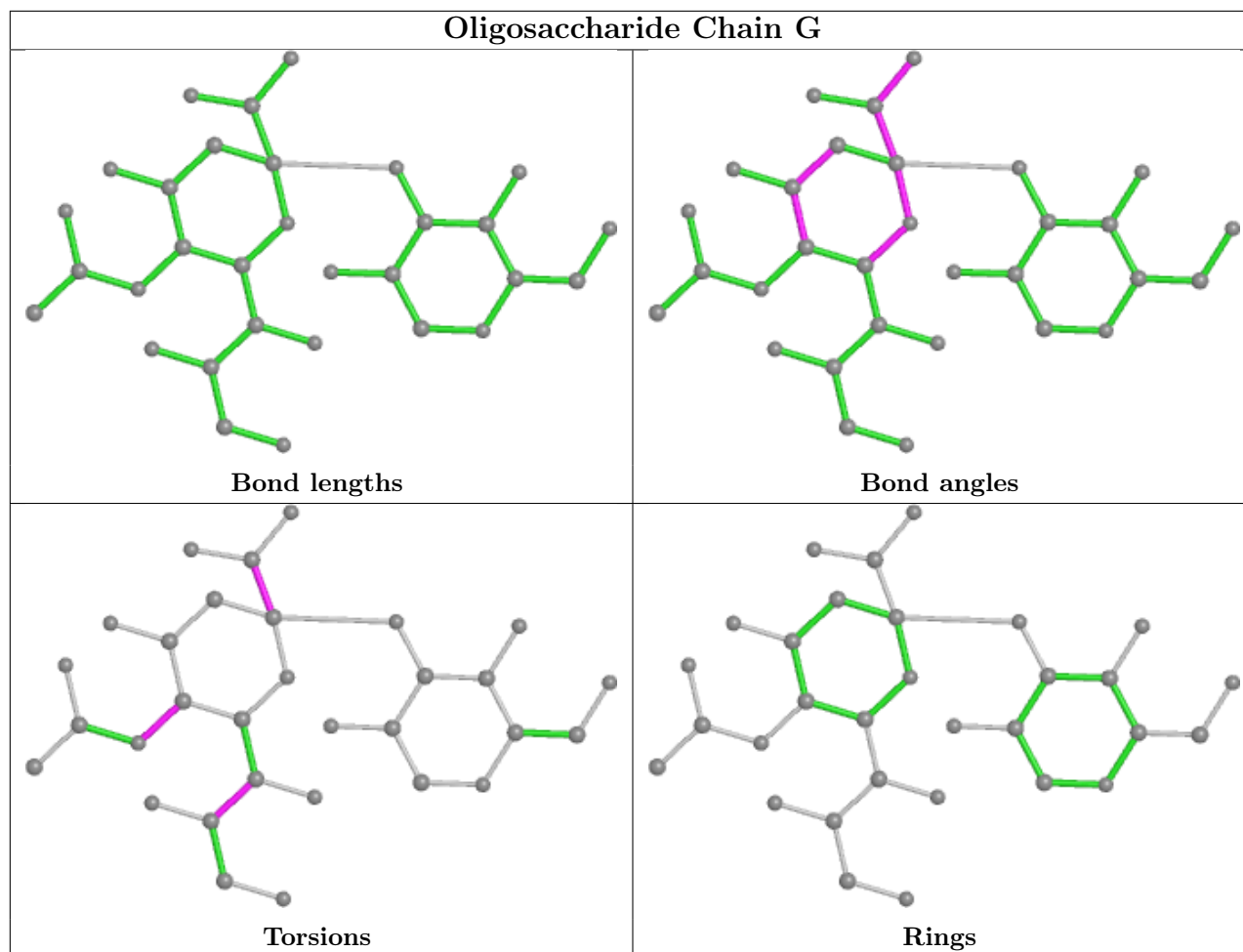
There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	4	SIA	1	0
4	F	2	NAG	1	0
3	E	1	NAG	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](#)

2 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	NAG	A	403	1	14,14,15	0.28	0	17,19,21	0.47	0
6	NAG	B	201	2	14,14,15	0.31	0	17,19,21	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	NAG	A	403	1	-	3/6/23/26	0/1/1/1
6	NAG	B	201	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	403	NAG	O5-C5-C6-O6
6	A	403	NAG	C4-C5-C6-O6
6	A	403	NAG	C3-C2-N2-C7

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	201	NAG	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	318/321 (99%)	-0.05	2 (0%) 89 91	29, 51, 85, 130	0
1	C	318/321 (99%)	1.18	68 (21%) 0 0	61, 108, 187, 237	0
2	B	169/183 (92%)	0.00	0 100 100	27, 45, 83, 152	0
2	D	169/183 (92%)	2.02	68 (40%) 0 0	79, 170, 209, 270	0
All	All	974/1008 (96%)	0.72	138 (14%) 2 1	27, 79, 190, 270	0

The worst 5 of 138 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	319	GLY	17.8
2	D	23	GLY	9.5
1	C	58	THR	9.1
2	D	36	ALA	8.3
1	C	327	PRO	8.3

### 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.

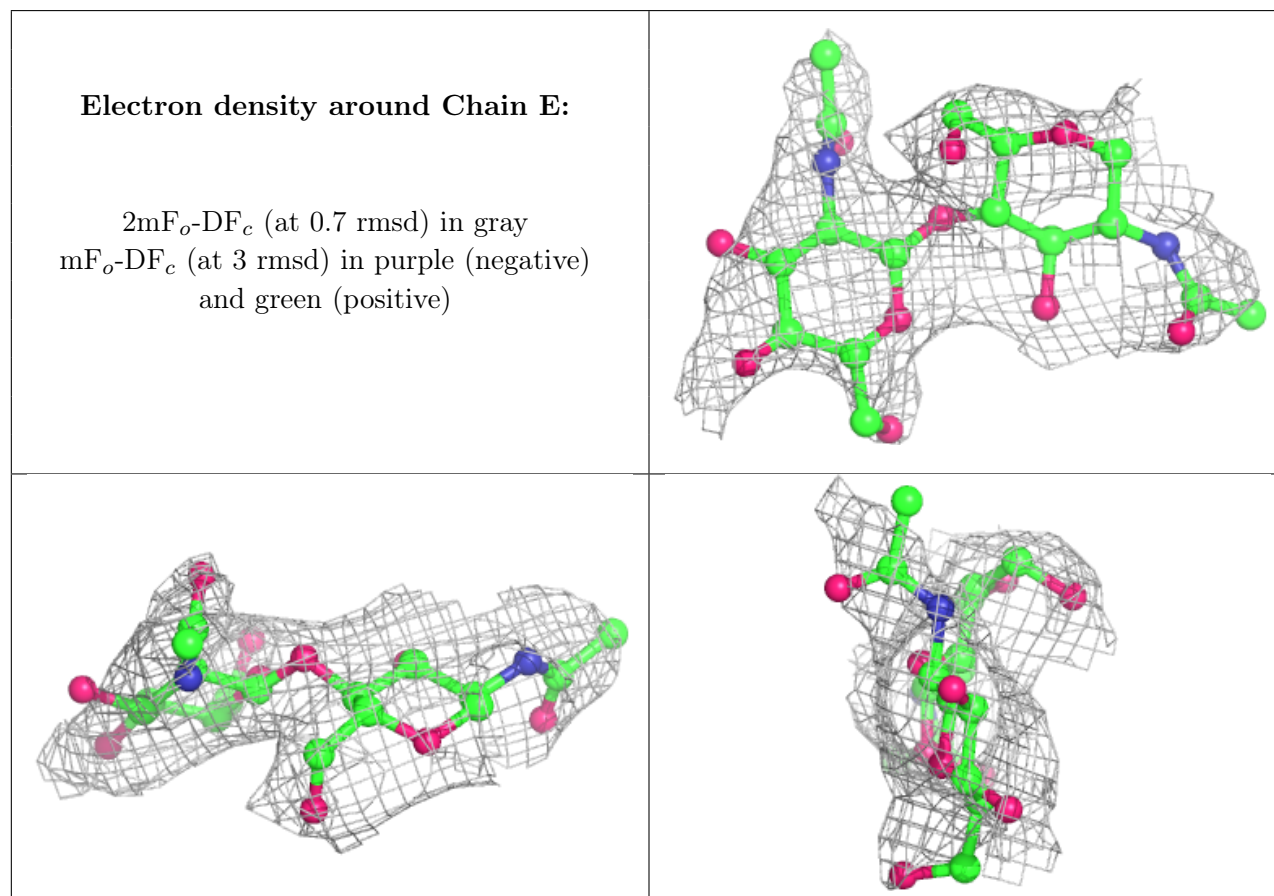
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NGA	F	1	15/15	0.62	0.38	169,179,192,193	0
3	NAG	E	2	14/15	0.81	0.32	118,124,128,129	0

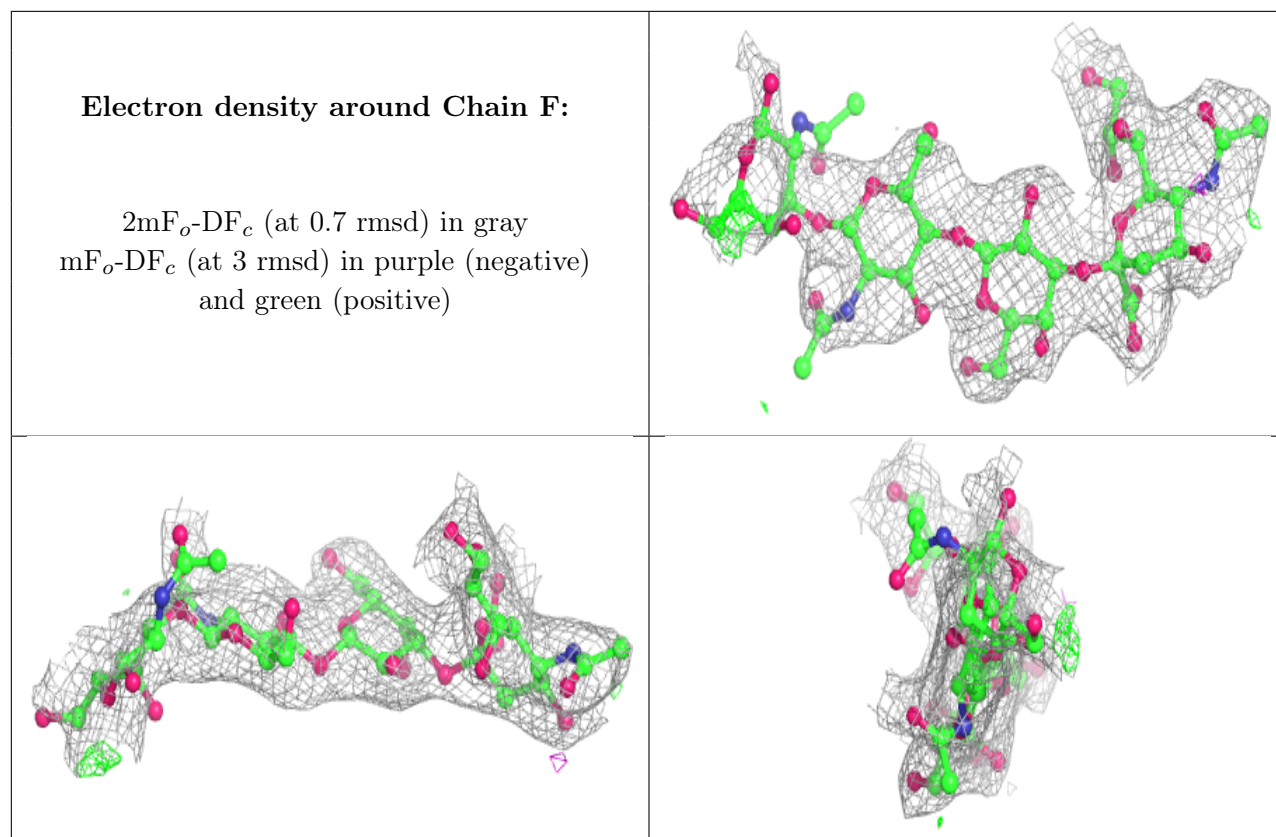
*Continued on next page...*

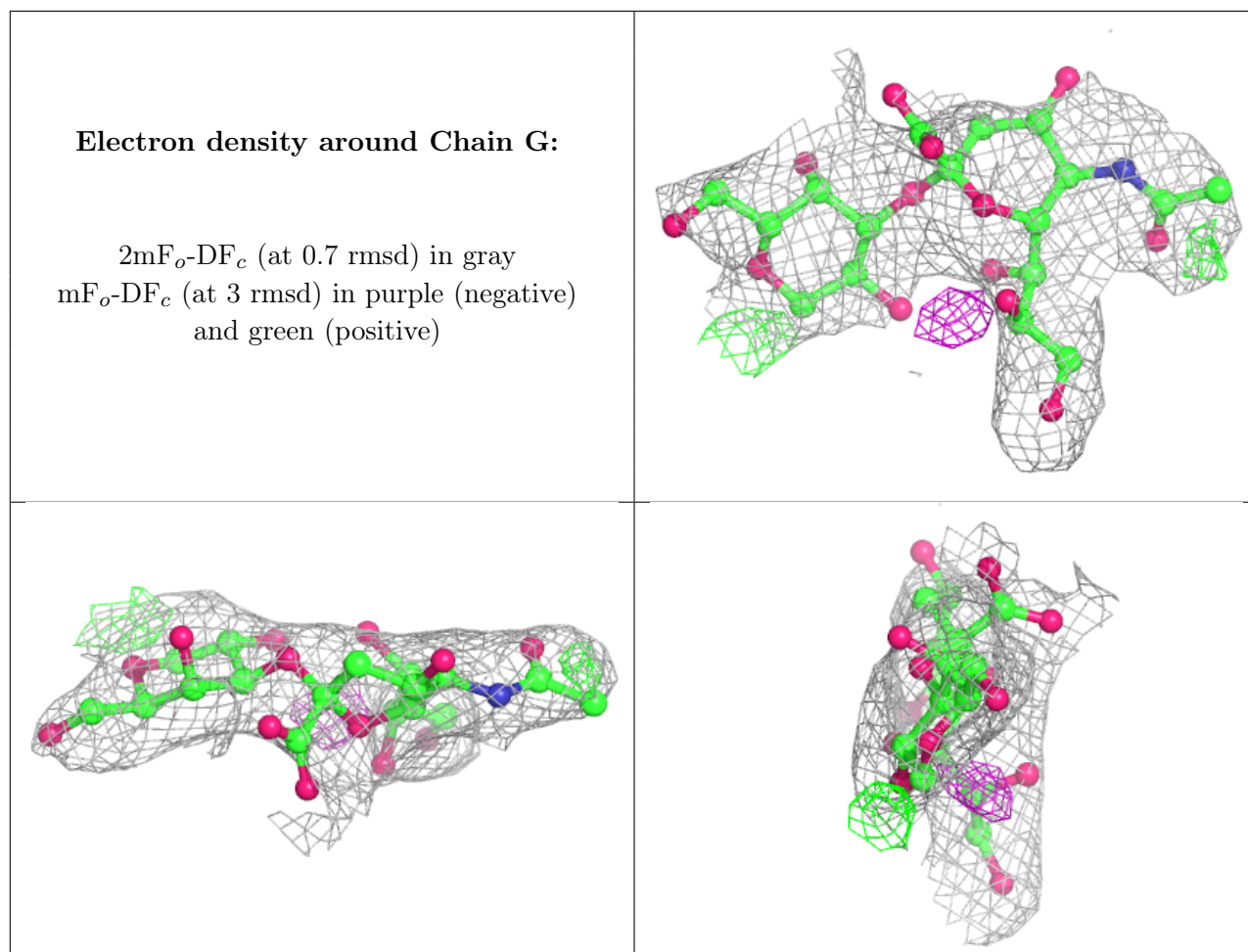
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	SIA	G	2	20/21	0.82	0.27	65,97,121,124	0
5	GAL	G	1	11/12	0.87	0.21	117,119,122,122	0
4	NAG	F	2	14/15	0.88	0.25	95,115,121,126	0
3	NAG	E	1	14/15	0.89	0.20	79,97,103,108	0
4	GAL	F	3	11/12	0.94	0.18	64,71,79,90	0
4	SIA	F	4	20/21	0.95	0.18	48,59,69,70	0

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







## 6.4 Ligands [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
6	NAG	A	403	14/15	0.67	0.21	127,137,145,147	0
6	NAG	B	201	14/15	0.91	0.20	69,80,86,91	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.