



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2020 – 06:18 PM BST

PDB ID : 6U7G  
Title : HCoV-229E RBD Class V in complex with human APN  
Authors : Tomlinson, A.; Li, Z.; Rini, J.M.  
Deposited on : 2019-09-02  
Resolution : 2.35 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

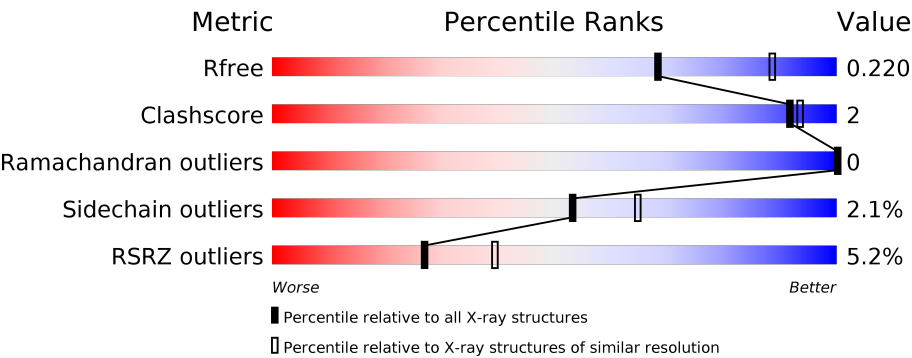
MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:  
*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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.



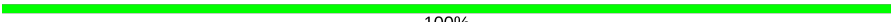

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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	906	<div><div>%</div><div><div></div><div>95%</div><div></div></div><div><div></div><div></div><div></div></div></div>
1	B	906	<div><div>%</div><div><div></div><div>94%</div><div></div></div><div><div></div><div></div><div>5%</div></div></div>
2	C	145	<div><div>21%</div><div><div></div><div>78%</div><div></div></div><div><div></div><div></div><div>18%</div></div></div>
2	D	145	<div><div>37%</div><div><div></div><div>66%</div><div></div></div><div><div></div><div>10%</div><div>23%</div></div></div>
3	E	2	<div><div></div><div><div></div><div>100%</div><div></div></div><div><div></div></div></div>
3	F	2	<div><div></div><div><div></div><div>50%</div><div></div></div><div><div></div><div>50%</div></div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	2	 100%
3	H	2	 50%50%

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 33302 atoms, of which 15945 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 Aminopeptidase N.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	898	Total	C	H	N	O	S	0	0	0
			14262	4623	7012	1223	1380	24			
1	B	899	Total	C	H	N	O	S	0	0	0
			14297	4634	7030	1225	1384	24			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	62	GLY	-	cloning artifact	UNP P15144
A	63	GLY	-	cloning artifact	UNP P15144
A	64	ARG	-	cloning artifact	UNP P15144
A	65	PRO	-	cloning artifact	UNP P15144
B	62	GLY	-	cloning artifact	UNP P15144
B	63	GLY	-	cloning artifact	UNP P15144
B	64	ARG	-	cloning artifact	UNP P15144
B	65	PRO	-	cloning artifact	UNP P15144

- Molecule 2 is a protein called Spike protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	119	Total	C	H	N	O	S	0	0	0
			1759	579	857	155	159	9			
2	D	111	Total	C	H	N	O	S	0	0	0
			1650	537	809	146	150	8			

There are 16 discrepancies between the modelled and reference sequences:

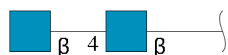
Chain	Residue	Modelled	Actual	Comment	Reference
C	289	GLY	-	cloning artifact	UNP H1AG31
C	290	GLY	-	cloning artifact	UNP H1AG31
C	291	ARG	-	cloning artifact	UNP H1AG31
C	292	PRO	-	cloning artifact	UNP H1AG31

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	300	MET	THR	conflict	UNP H1AG31
C	412	THR	SER	conflict	UNP H1AG31
C	422	GLY	VAL	conflict	UNP H1AG31
C	428	GLN	LYS	conflict	UNP H1AG31
D	289	GLY	-	cloning artifact	UNP H1AG31
D	290	GLY	-	cloning artifact	UNP H1AG31
D	291	ARG	-	cloning artifact	UNP H1AG31
D	292	PRO	-	cloning artifact	UNP H1AG31
D	300	MET	THR	conflict	UNP H1AG31
D	412	THR	SER	conflict	UNP H1AG31
D	422	GLY	VAL	conflict	UNP H1AG31
D	428	GLN	LYS	conflict	UNP H1AG31

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



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	2	Total	C	H	N	O	0	0	0
			56	16	28	2	10			
3	F	2	Total	C	H	N	O	0	0	0
			56	16	28	2	10			
3	G	2	Total	C	H	N	O	0	0	0
			56	16	28	2	10			
3	H	2	Total	C	N	O	0	0	0	
			28	16	2	10				

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Zn	0	0
			1	1		
4	A	1	Total	Zn	0	0
			1	1		

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



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
5	A	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
5	A	1	Total	C	N	O		0	0
			14	8	1	5			
5	A	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
5	A	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
5	B	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
5	B	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
5	B	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
5	B	1	Total	C	N	O		0	0
			14	8	1	5			
5	C	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
5	C	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
5	D	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
5	D	1	Total	C	H	N	O	0	0
			28	8	14	1	5		

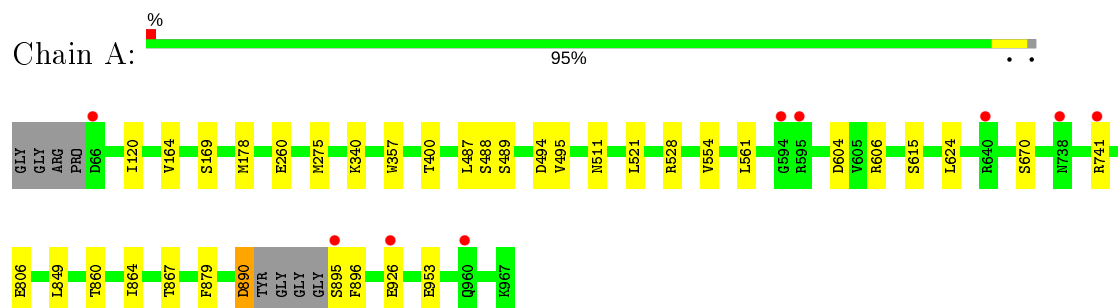
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	403	Total 403	O 403	0	0
6	B	350	Total 350	O 350	0	0
6	C	36	Total 36	O 36	0	0
6	D	12	Total 12	O 12	0	0

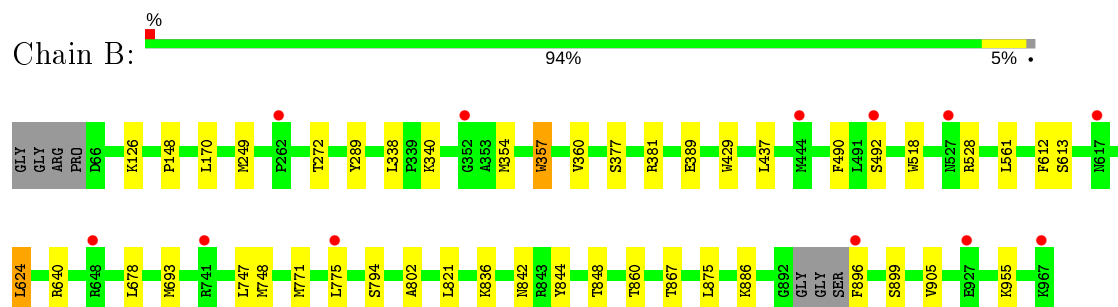
### 3 Residue-property plots [i](#)

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.

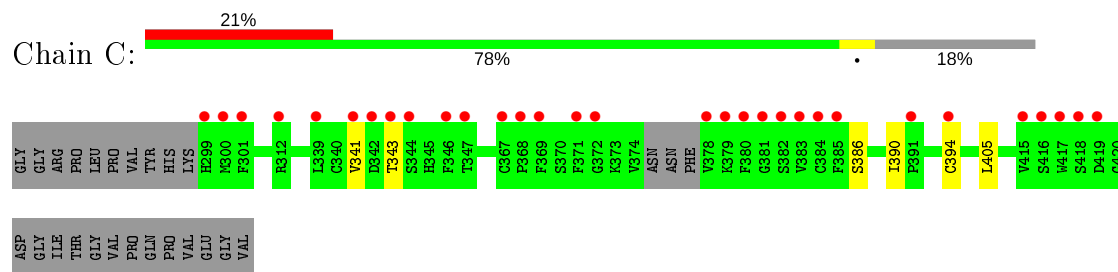
#### • Molecule 1: Aminopeptidase N



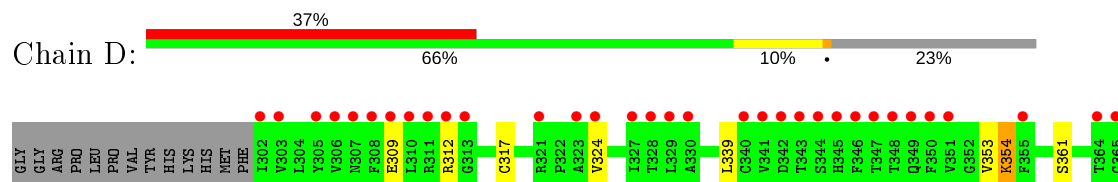
#### • Molecule 1: Aminopeptidase N



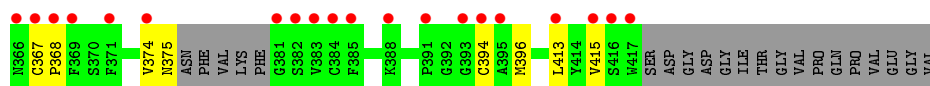
#### • Molecule 2: Spike protein



#### • Molecule 2: Spike protein







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

Chain E:  100%



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

Chain F:  50% 50%



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

Chain G:  100%



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

Chain H:  50% 50%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.56 Å 98.68 Å 147.53 Å 90.00° 104.60° 90.00°	Depositor
Resolution (Å)	45.61 – 2.35 49.59 – 2.35	Depositor EDS
% Data completeness (in resolution range)	96.0 (45.61-2.35) 96.0 (49.59-2.35)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.67 (at 2.34 Å)	Xtriage
Refinement program	PHENIX 1.16rc1_3535	Depositor
R, $R_{free}$	0.179 , 0.218 0.181 , 0.220	Depositor DCC
$R_{free}$ test set	5552 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.0	Xtriage
Anisotropy	0.109	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 43.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	33302	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.93% 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: ZN, NAG

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	0.31	0/7436	0.45	0/10127
1	B	0.32	0/7454	0.46	0/10151
2	C	0.30	0/925	0.51	0/1257
2	D	0.32	0/861	0.52	0/1171
All	All	0.31	0/16676	0.46	0/22706

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	7250	7012	7011	20	0
1	B	7267	7030	7028	19	0
2	C	902	857	857	3	0
2	D	841	809	811	9	0
3	E	28	28	25	0	0
3	F	28	28	25	2	0
3	G	28	28	25	0	0
3	H	28	0	24	1	0
4	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1	0	0	0	0
5	A	70	55	65	1	0
5	B	56	42	52	1	0
5	C	28	28	26	0	0
5	D	28	28	26	0	0
6	A	403	0	0	2	0
6	B	350	0	0	0	0
6	C	36	0	0	0	0
6	D	12	0	0	0	0
All	All	17357	15945	15975	51	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 2.

All (51) 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:624:LEU:HD21	3:F:1:NAG:H82	1.65	0.78
5:A:5010:NAG:H82	5:A:5010:NAG:O3	1.86	0.74
2:D:396:MET:HE2	2:D:413:LEU:HD12	1.73	0.70
1:B:747:LEU:HD23	1:B:748:MET:CE	2.28	0.63
2:D:309:GLU:HB2	2:D:324:VAL:HG22	1.84	0.60
1:B:842:ASN:HA	1:B:875:LEU:HD21	1.83	0.59
1:A:554:VAL:HG22	1:A:561:LEU:HD22	1.86	0.58
1:A:164:VAL:HG21	1:A:178:MET:HE2	1.85	0.56
1:A:120:ILE:HG21	1:A:178:MET:HE3	1.89	0.55
1:A:890:ASP:N	1:A:890:ASP:OD1	2.39	0.55
2:C:341:VAL:HG22	2:C:343:THR:OG1	2.06	0.55
2:D:394:CYS:SG	2:D:415:VAL:HB	2.47	0.55
1:B:624:LEU:HD22	3:H:1:NAG:H82	1.90	0.52
1:B:354:MET:O	1:B:360:VAL:HG13	2.10	0.52
1:A:554:VAL:HG22	1:A:561:LEU:CD2	2.40	0.51
2:D:367:CYS:CB	2:D:394:CYS:SG	2.99	0.50
2:D:367:CYS:SG	2:D:368:PRO:HD2	2.53	0.48
2:C:386:SER:HB3	2:C:390:ILE:HD11	1.95	0.48
1:B:844:TYR:O	1:B:848:THR:HG23	2.14	0.47
1:B:561:LEU:HD11	1:B:612:PHE:CE2	2.49	0.47
1:A:860:THR:HG22	1:A:896:PHE:CE1	2.51	0.46
1:A:864:ILE:O	1:A:867:THR:HG22	2.15	0.46
1:B:867:THR:HG21	1:B:905:VAL:HA	1.99	0.45
1:B:437:LEU:HD11	1:B:693:MET:HG2	1.98	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:353:VAL:HG22	2:D:354:LYS:N	2.32	0.45
1:B:148:PRO:HD2	1:B:170:LEU:HD23	1.98	0.45
1:B:771:MET:SD	1:B:775:LEU:HD12	2.57	0.45
1:A:494:ASP:OD1	1:A:528:ARG:NH2	2.48	0.44
1:B:747:LEU:HD23	1:B:748:MET:HE3	1.99	0.44
1:B:249:MET:HG2	1:B:272:THR:HG22	1.99	0.44
1:A:806:GLU:OE1	6:A:5101:HOH:O	2.20	0.44
1:A:864:ILE:HA	1:A:867:THR:HG22	2.00	0.44
1:A:860:THR:HG23	6:A:5212:HOH:O	2.17	0.44
1:B:678:LEU:HB2	5:B:5009:NAG:H82	2.01	0.43
1:B:860:THR:HG23	1:B:896:PHE:CD1	2.53	0.43
1:A:487:LEU:HD21	1:A:521:LEU:CD1	2.49	0.43
1:A:488:SER:O	1:A:488:SER:OG	2.37	0.43
2:D:374:VAL:O	2:D:375:ASN:HB3	2.19	0.42
1:B:360:VAL:HG11	1:B:389:GLU:HB3	2.02	0.42
1:B:429:TRP:CD2	1:B:747:LEU:HD22	2.54	0.42
1:A:495:VAL:HG22	1:A:528:ARG:HG3	2.02	0.42
2:D:367:CYS:HB2	2:D:394:CYS:SG	2.60	0.42
1:A:120:ILE:HG21	1:A:178:MET:CE	2.50	0.41
1:A:849:LEU:HD11	1:A:879:PHE:CZ	2.56	0.41
1:A:860:THR:HG22	1:A:896:PHE:CD1	2.56	0.41
1:A:624:LEU:CD2	3:F:1:NAG:H82	2.42	0.41
1:B:338:LEU:HD11	1:B:357:TRP:CD1	2.55	0.41
1:A:400:THR:O	1:A:511:ASN:HA	2.20	0.40
1:B:802:ALA:O	1:B:836:LYS:HD2	2.21	0.40
2:C:405:LEU:HD12	2:C:405:LEU:HA	1.90	0.40
1:B:289:TYR:CZ	2:D:317:CYS:HB2	2.56	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	894/906 (99%)	875 (98%)	19 (2%)	0	100	100
1	B	895/906 (99%)	876 (98%)	19 (2%)	0	100	100
2	C	115/145 (79%)	108 (94%)	7 (6%)	0	100	100
2	D	107/145 (74%)	104 (97%)	3 (3%)	0	100	100
All	All	2011/2102 (96%)	1963 (98%)	48 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	797/802 (99%)	782 (98%)	15 (2%)	57	68
1	B	799/802 (100%)	782 (98%)	17 (2%)	53	65
2	C	97/123 (79%)	96 (99%)	1 (1%)	76	85
2	D	92/123 (75%)	88 (96%)	4 (4%)	29	35
All	All	1785/1850 (96%)	1748 (98%)	37 (2%)	53	65

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	169	SER
1	A	260	GLU
1	A	275	MET
1	A	340	LYS
1	A	357	TRP
1	A	489	SER
1	A	604	ASP
1	A	606	ARG
1	A	615	SER
1	A	670	SER
1	A	741	ARG
1	A	890	ASP
1	A	895	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	926	GLU
1	A	953	GLU
1	B	126	LYS
1	B	340	LYS
1	B	357	TRP
1	B	377	SER
1	B	381	ARG
1	B	490	PHE
1	B	492	SER
1	B	518	TRP
1	B	528	ARG
1	B	613	SER
1	B	624	LEU
1	B	640	ARG
1	B	794	SER
1	B	821	LEU
1	B	886	LYS
1	B	899	SER
1	B	955	LYS
2	C	394	CYS
2	D	312	ARG
2	D	339	LEU
2	D	354	LYS
2	D	361	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	146	GLN
1	B	175	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

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	1,3	14,14,15	0.21	0	17,19,21	0.43	0
3	NAG	E	2	3	14,14,15	0.19	0	17,19,21	0.44	0
3	NAG	F	1	1,3	14,14,15	0.17	0	17,19,21	0.47	0
3	NAG	F	2	3	14,14,15	0.22	0	17,19,21	0.39	0
3	NAG	G	1	1,3	14,14,15	0.24	0	17,19,21	0.38	0
3	NAG	G	2	3	14,14,15	0.24	0	17,19,21	0.37	0
3	NAG	H	1	1,3	14,14,15	2.21	4 (28%)	17,19,21	2.03	7 (41%)
3	NAG	H	2	3	14,14,15	1.24	1 (7%)	17,19,21	1.73	6 (35%)

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	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	2/6/23/26	0/1/1/1
3	NAG	F	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	2/6/23/26	0/1/1/1
3	NAG	G	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	G	2	3	-	2/6/23/26	0/1/1/1
3	NAG	H	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	H	2	3	-	2/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	H	1	NAG	O5-C1	-4.74	1.36	1.43
3	H	1	NAG	C1-C2	3.58	1.57	1.52
3	H	2	NAG	O5-C1	-3.34	1.38	1.43

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	H	1	NAG	O5-C5	-3.22	1.36	1.43
3	H	1	NAG	O7-C7	-2.19	1.18	1.23

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	1	NAG	O3-C3-C4	-4.15	100.76	110.35
3	H	1	NAG	O4-C4-C3	-3.40	102.50	110.35
3	H	2	NAG	C2-N2-C7	-3.21	118.33	122.90
3	H	2	NAG	O5-C1-C2	-3.07	106.44	111.29
3	H	2	NAG	C4-C3-C2	-2.95	106.70	111.02
3	H	1	NAG	O3-C3-C2	-2.58	104.12	109.47
3	H	1	NAG	C2-N2-C7	-2.53	119.29	122.90
3	H	1	NAG	C4-C3-C2	-2.44	107.44	111.02
3	H	2	NAG	C3-C4-C5	-2.42	105.91	110.24
3	H	1	NAG	O5-C5-C6	-2.31	103.58	107.20
3	H	1	NAG	C1-C2-N2	-2.29	106.58	110.49
3	H	2	NAG	O5-C5-C4	-2.16	105.56	110.83
3	H	2	NAG	C1-O5-C5	2.04	114.96	112.19

There are no chirality outliers.

All (10) torsion outliers are listed below:

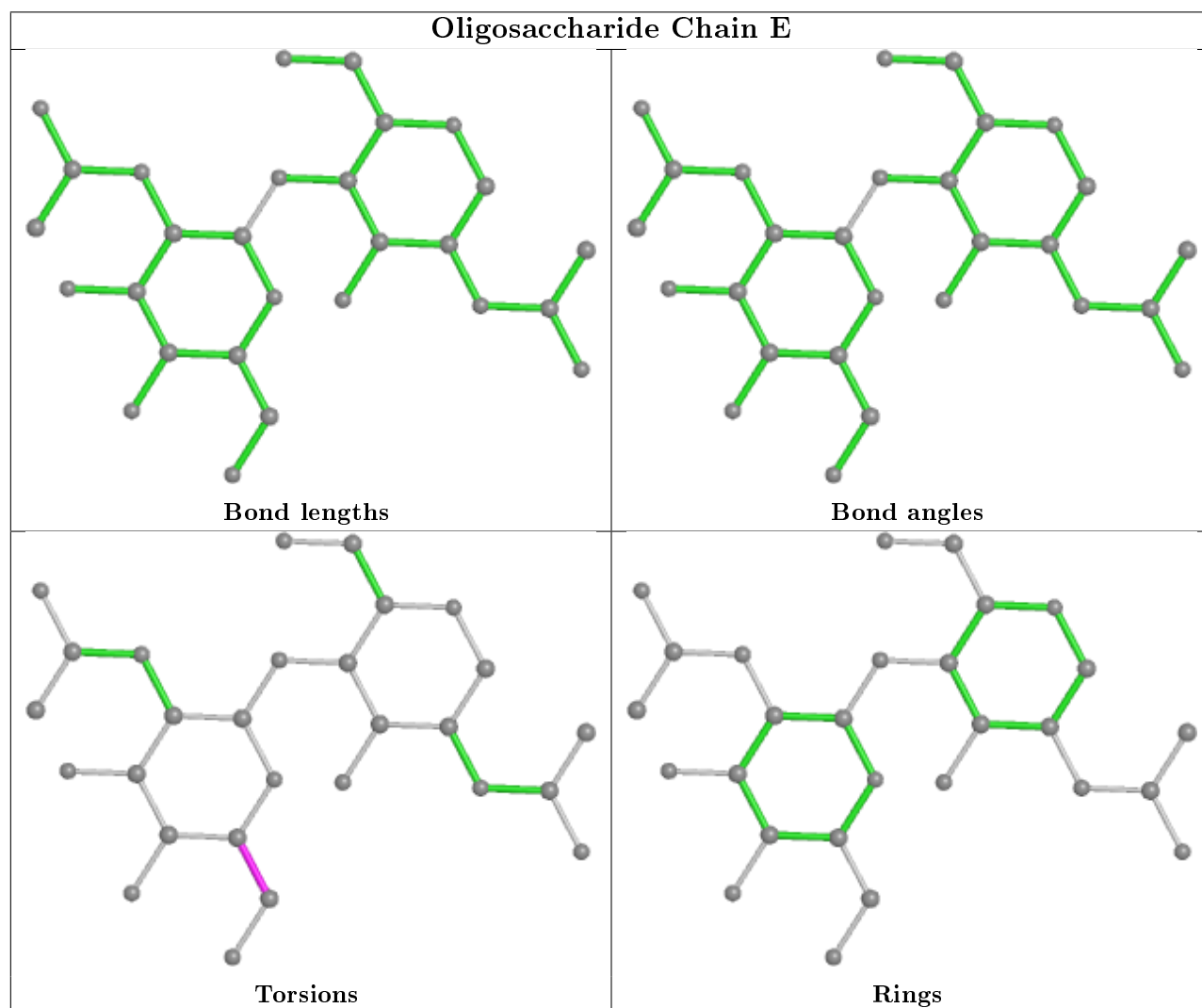
Mol	Chain	Res	Type	Atoms
3	H	2	NAG	O5-C5-C6-O6
3	H	2	NAG	C4-C5-C6-O6
3	F	2	NAG	O5-C5-C6-O6
3	H	1	NAG	C4-C5-C6-O6
3	H	1	NAG	O5-C5-C6-O6
3	E	2	NAG	C4-C5-C6-O6
3	F	2	NAG	C4-C5-C6-O6
3	G	2	NAG	C4-C5-C6-O6
3	G	2	NAG	O5-C5-C6-O6
3	E	2	NAG	O5-C5-C6-O6

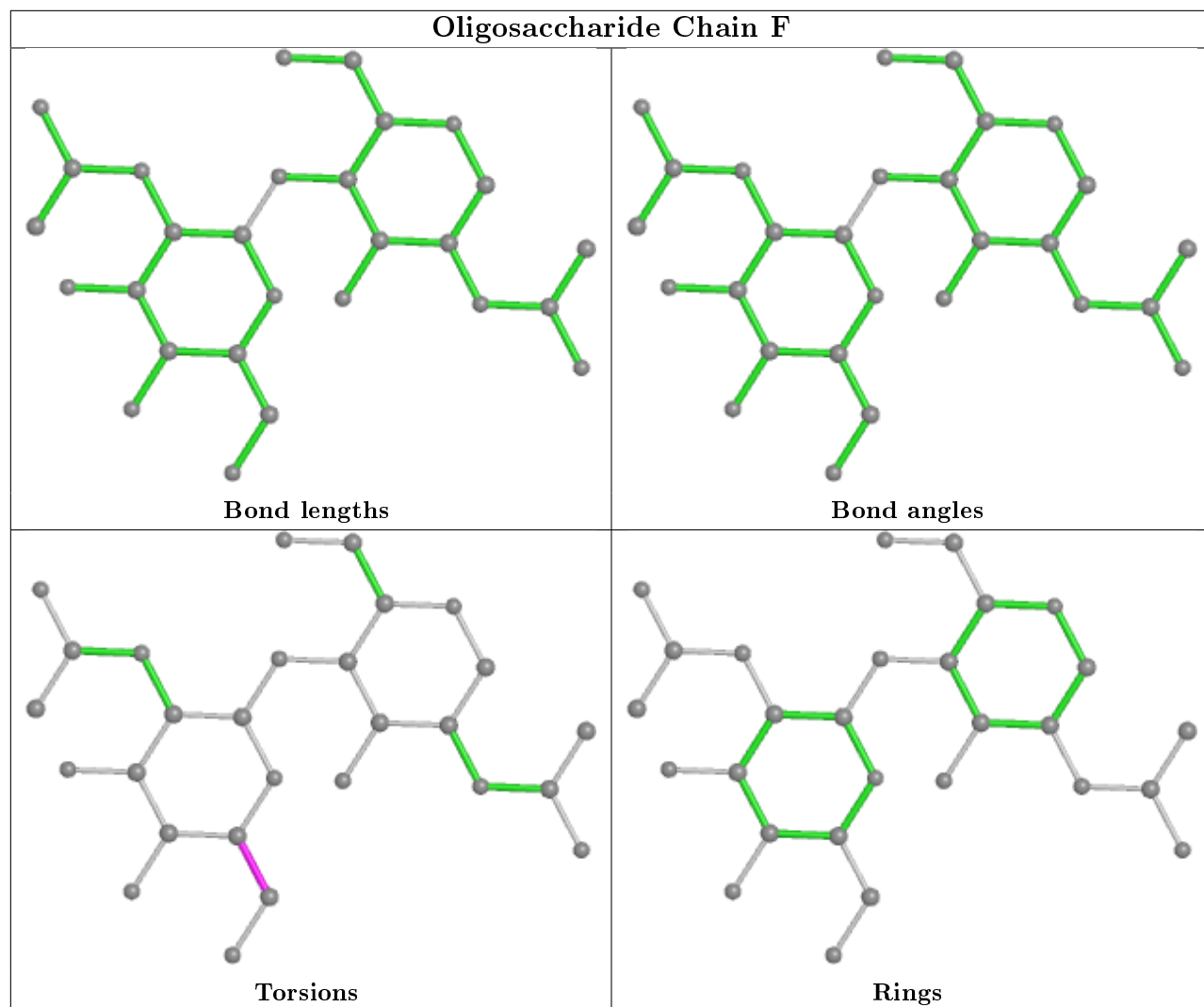
There are no ring outliers.

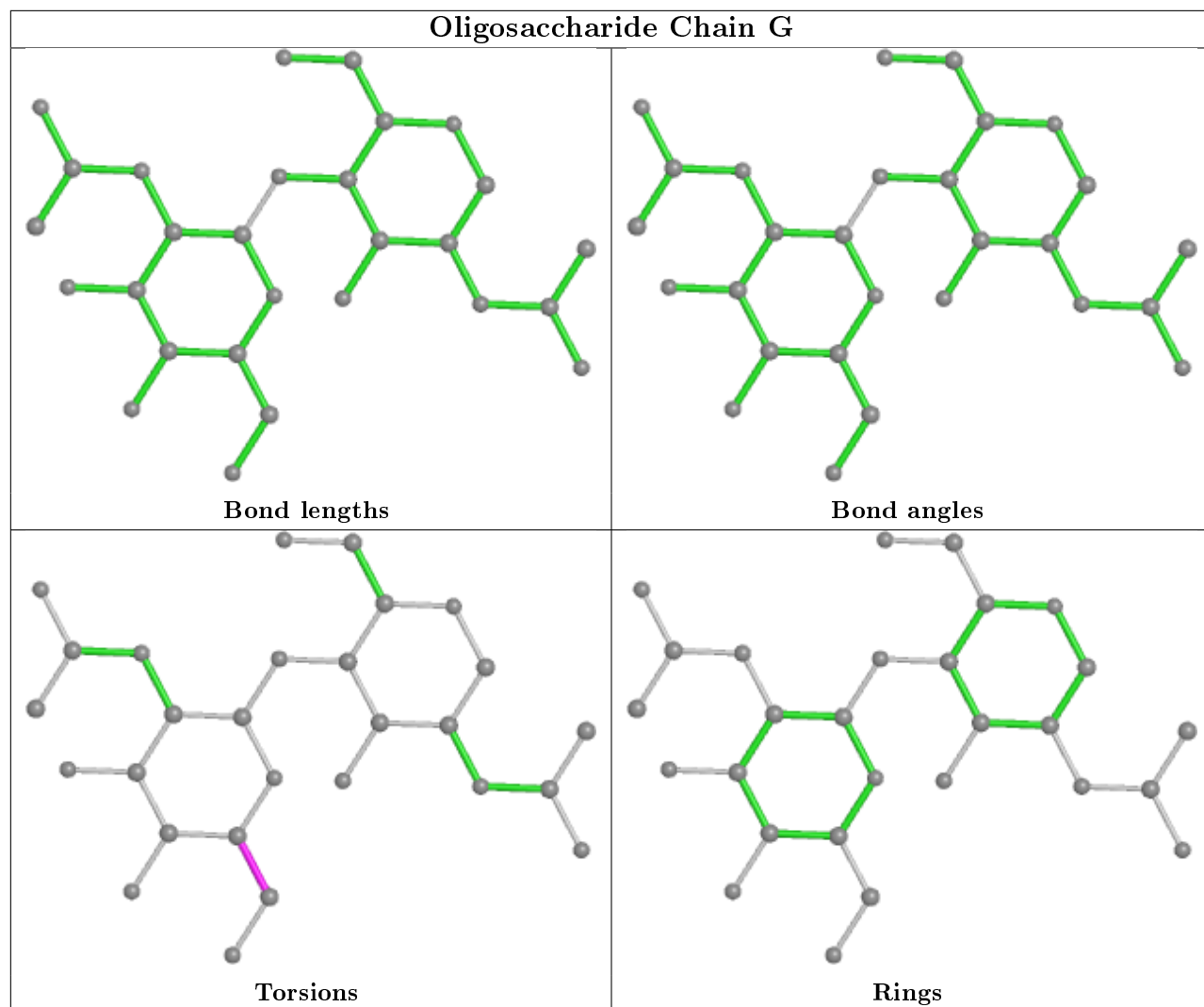
2 monomers are involved in 3 short contacts:

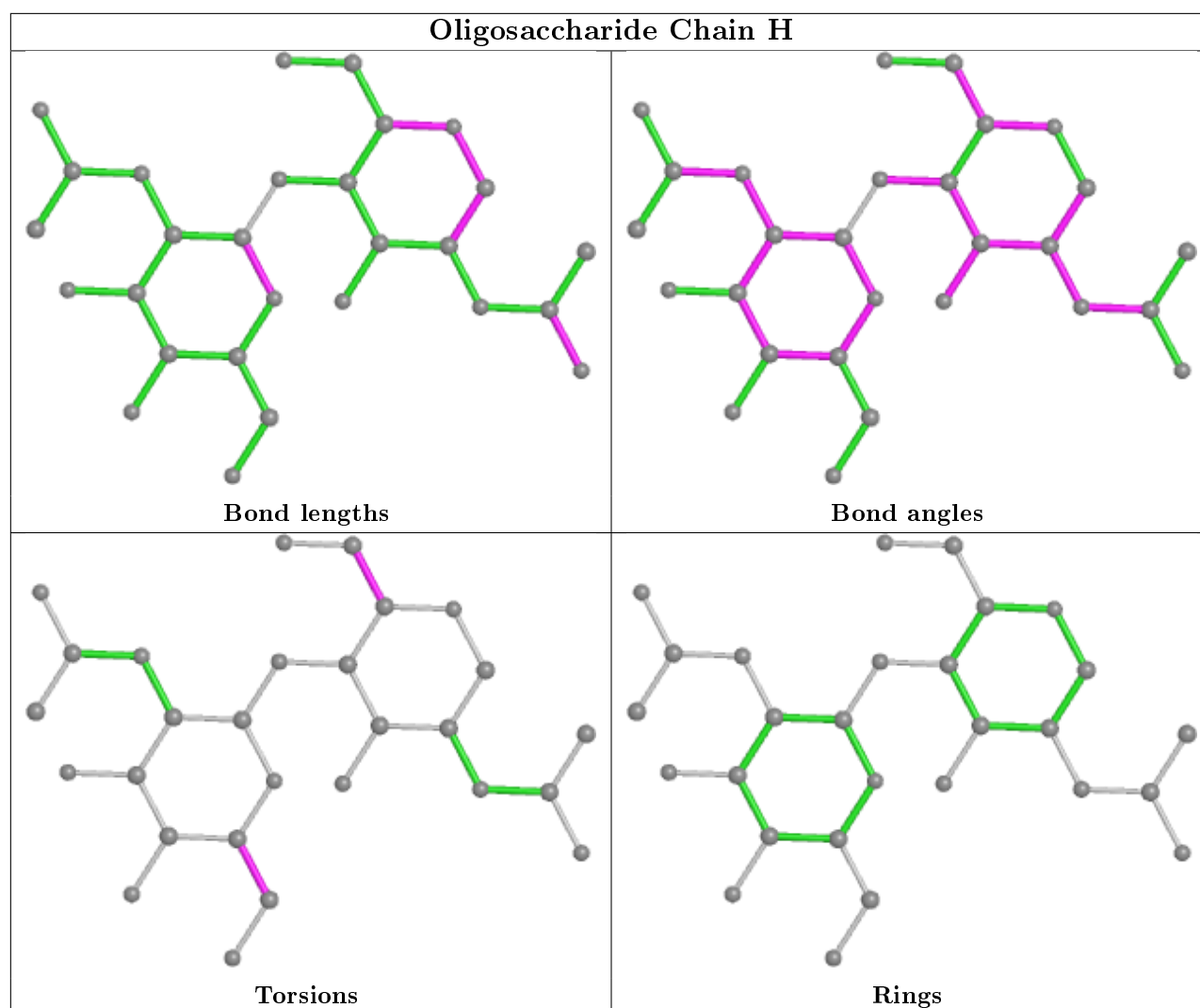
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	1	NAG	2	0
3	H	1	NAG	1	0

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](#)

Of 15 ligands modelled in this entry, 2 are monoatomic - leaving 13 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	NAG	A	5006	1	14,14,15	1.74	5 (35%)	17,19,21	1.72	4 (23%)
5	NAG	C	2001	2	14,14,15	0.32	0	17,19,21	0.47	0
5	NAG	A	5005	1	14,14,15	0.29	0	17,19,21	0.38	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	B	5005	1	14,14,15	0.25	0	17,19,21	0.38	0
5	NAG	C	2002	2	14,14,15	0.34	0	17,19,21	0.52	0
5	NAG	B	5006	1	14,14,15	0.17	0	17,19,21	0.45	0
5	NAG	A	5002	1	14,14,15	0.27	0	17,19,21	0.38	0
5	NAG	B	5002	1	14,14,15	0.21	0	17,19,21	0.41	0
5	NAG	D	502	2	14,14,15	0.59	0	17,19,21	1.47	1 (5%)
5	NAG	A	5007	1	14,14,15	0.21	0	17,19,21	0.40	0
5	NAG	A	5010	1	14,14,15	0.28	0	17,19,21	0.58	0
5	NAG	D	501	2	14,14,15	0.28	0	17,19,21	0.52	0
5	NAG	B	5009	1	14,14,15	1.44	2 (14%)	17,19,21	2.03	4 (23%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	A	5006	1	-	0/6/23/26	0/1/1/1
5	NAG	C	2001	2	-	0/6/23/26	0/1/1/1
5	NAG	A	5005	1	-	0/6/23/26	0/1/1/1
5	NAG	B	5005	1	-	0/6/23/26	0/1/1/1
5	NAG	C	2002	2	-	0/6/23/26	0/1/1/1
5	NAG	B	5006	1	-	0/6/23/26	0/1/1/1
5	NAG	A	5002	1	-	2/6/23/26	0/1/1/1
5	NAG	B	5002	1	-	2/6/23/26	0/1/1/1
5	NAG	D	502	2	-	2/6/23/26	0/1/1/1
5	NAG	A	5007	1	-	0/6/23/26	0/1/1/1
5	NAG	A	5010	1	-	2/6/23/26	0/1/1/1
5	NAG	D	501	2	-	2/6/23/26	0/1/1/1
5	NAG	B	5009	1	-	0/6/23/26	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	5006	NAG	O5-C5	-2.97	1.37	1.43
5	A	5006	NAG	O7-C7	-2.83	1.16	1.23
5	B	5009	NAG	O5-C1	-2.55	1.39	1.43
5	B	5009	NAG	O5-C5	-2.23	1.38	1.43
5	A	5006	NAG	C2-N2	-2.18	1.42	1.46
5	A	5006	NAG	O5-C1	-2.15	1.40	1.43
5	A	5006	NAG	C4-C5	-2.01	1.48	1.53

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	502	NAG	C1-O5-C5	5.63	119.82	112.19
5	B	5009	NAG	O5-C1-C2	-5.42	102.73	111.29
5	B	5009	NAG	C1-C2-N2	3.27	116.08	110.49
5	A	5006	NAG	C3-C4-C5	-3.20	104.54	110.24
5	B	5009	NAG	C2-N2-C7	-3.07	118.53	122.90
5	A	5006	NAG	C8-C7-N2	2.93	121.06	116.10
5	A	5006	NAG	C1-O5-C5	2.61	115.73	112.19
5	A	5006	NAG	O4-C4-C5	-2.32	103.53	109.30
5	B	5009	NAG	O6-C6-C5	-2.23	103.63	111.29

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	D	502	NAG	O5-C5-C6-O6
5	A	5010	NAG	C8-C7-N2-C2
5	A	5010	NAG	O7-C7-N2-C2
5	D	502	NAG	C4-C5-C6-O6
5	A	5002	NAG	O5-C5-C6-O6
5	A	5002	NAG	C4-C5-C6-O6
5	B	5002	NAG	C4-C5-C6-O6
5	B	5002	NAG	O5-C5-C6-O6
5	D	501	NAG	C4-C5-C6-O6
5	D	501	NAG	O5-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	5010	NAG	1	0
5	B	5009	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 ⓘ

### 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	898/906 (99%)	0.17	9 (1%) 82 88	17, 33, 58, 81	0
1	B	899/906 (99%)	0.17	12 (1%) 77 84	18, 34, 58, 80	0
2	C	119/145 (82%)	1.24	31 (26%) 0 0	28, 51, 98, 116	0
2	D	111/145 (76%)	2.11	53 (47%) 0 0	39, 64, 96, 114	0
All	All	2027/2102 (96%)	0.34	105 (5%) 27 39	17, 35, 69, 116	0

All (105) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	374	VAL	6.2
2	C	343	THR	6.1
2	C	341	VAL	6.0
2	D	341	VAL	6.0
2	C	382	SER	5.9
2	D	417	TRP	5.9
2	C	417	TRP	5.8
2	C	300	MET	5.8
2	D	343	THR	5.6
2	C	344	SER	5.2
2	D	347	THR	5.2
2	D	346	PHE	5.2
2	C	342	ASP	5.1
2	D	382	SER	5.0
2	D	415	VAL	4.9
2	D	366	ASN	4.8
2	D	345	HIS	4.7
2	C	419	ASP	4.6
2	D	368	PRO	4.6
2	C	369	PHE	4.5
2	D	369	PHE	4.5

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	303	VAL	4.5
2	D	416	SER	4.4
2	D	344	SER	4.1
2	C	379	LYS	4.0
2	D	383	VAL	3.9
2	D	310	LEU	3.9
2	C	299	HIS	3.8
2	C	418	SER	3.7
2	C	415	VAL	3.7
1	B	352	GLY	3.6
1	A	738	ASN	3.6
2	D	312	ARG	3.6
2	C	391	PRO	3.5
2	D	305	TYR	3.5
2	D	384	CYS	3.5
2	D	342	ASP	3.4
2	C	368	PRO	3.4
2	D	367	CYS	3.4
2	C	339	LEU	3.3
2	C	301	PHE	3.3
2	D	327	ILE	3.3
2	D	348	THR	3.2
2	C	381	GLY	3.2
2	D	323	ALA	3.2
1	A	594	GLY	3.1
1	B	967	LYS	3.1
2	D	365	GLY	3.1
2	D	308	PHE	3.1
2	C	394	CYS	3.1
2	C	378	VAL	3.0
2	C	383	VAL	3.0
2	D	388	LYS	3.0
2	D	329	LEU	2.9
2	C	385	PHE	2.9
2	D	371	PHE	2.9
2	D	381	GLY	2.9
2	C	372	GLY	2.9
2	D	302	ILE	2.8
1	A	595	ARG	2.8
1	A	741	ARG	2.8
2	C	380	PHE	2.8
2	D	385	PHE	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	311	ARG	2.8
2	D	349	GLN	2.7
1	B	927	GLU	2.7
1	B	617	ASN	2.7
2	C	346	PHE	2.7
2	D	306	VAL	2.7
2	D	355	PHE	2.6
2	D	393	GLY	2.6
2	D	340	CYS	2.6
2	D	350	PHE	2.6
2	D	391	PRO	2.6
2	D	330	ALA	2.6
2	D	395	ALA	2.6
2	C	347	THR	2.5
1	B	896	PHE	2.5
2	D	394	CYS	2.5
2	D	307	ASN	2.4
1	A	895	SER	2.4
2	D	313	GLY	2.4
2	D	309	GLU	2.4
2	C	312	ARG	2.4
1	A	926	GLU	2.4
2	C	416	SER	2.3
2	D	351	VAL	2.3
2	C	367	CYS	2.3
2	D	413	LEU	2.2
1	B	492	SER	2.2
1	A	66	ASP	2.2
2	D	328	THR	2.1
1	B	741	ARG	2.1
2	C	371	PHE	2.1
1	A	960	GLN	2.1
1	B	262	PRO	2.1
1	B	444	MET	2.0
2	D	324	VAL	2.0
2	C	384	CYS	2.0
1	A	640	ARG	2.0
1	B	648	ARG	2.0
1	B	527	ASN	2.0
1	B	775	LEU	2.0
2	D	321	ARG	2.0
2	D	364	THR	2.0

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

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

## 6.3 Carbohydrates ⓘ

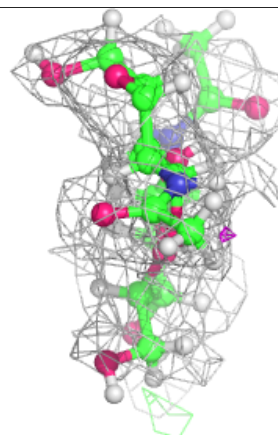
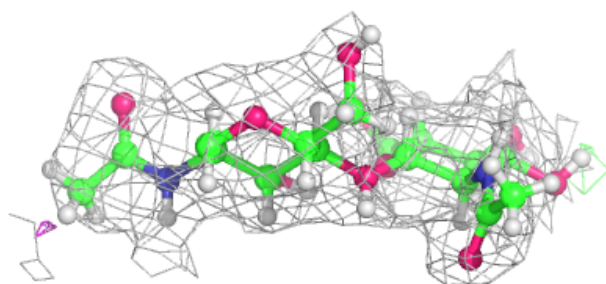
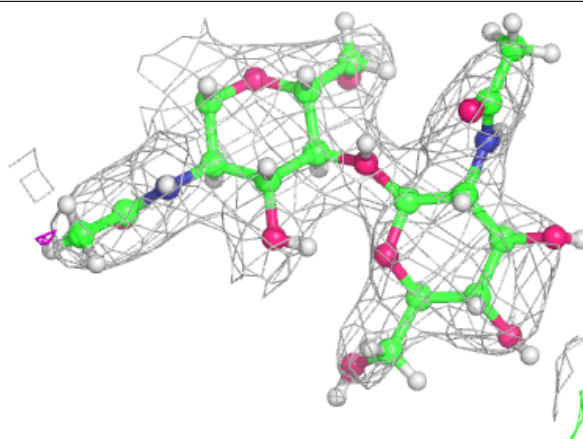
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
3	NAG	F	2	14/15	0.81	0.28	80,96,113,118	0
3	NAG	H	2	14/15	0.86	0.19	69,76,87,88	0
3	NAG	E	2	14/15	0.87	0.25	61,74,88,92	0
3	NAG	G	2	14/15	0.88	0.27	67,81,93,99	0
3	NAG	F	1	14/15	0.94	0.16	50,66,78,84	0
3	NAG	E	1	14/15	0.95	0.15	37,48,61,70	0
3	NAG	G	1	14/15	0.95	0.12	39,49,60,69	0
3	NAG	H	1	14/15	0.96	0.09	42,47,57,68	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.

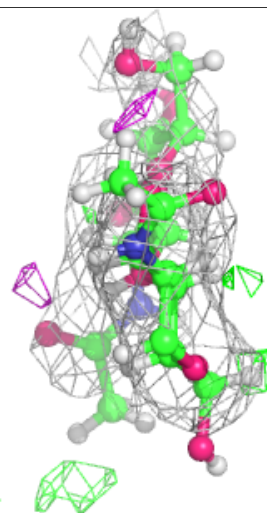
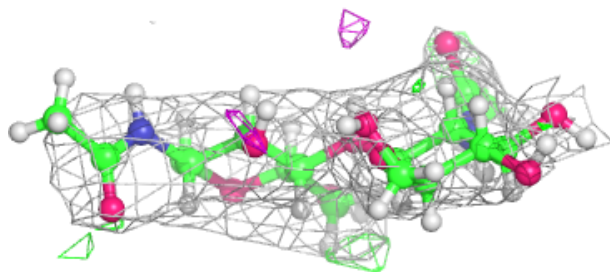
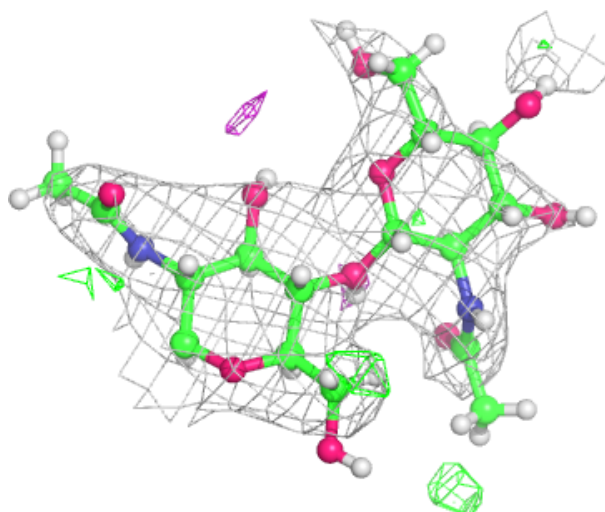
**Electron density around Chain E:**

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



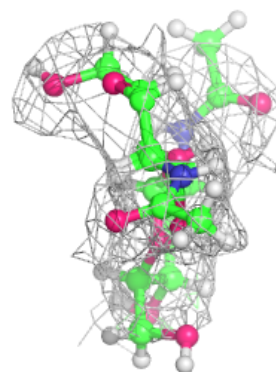
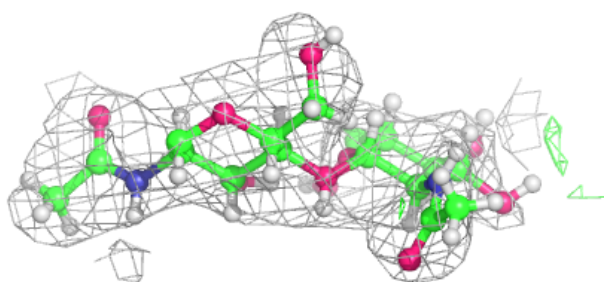
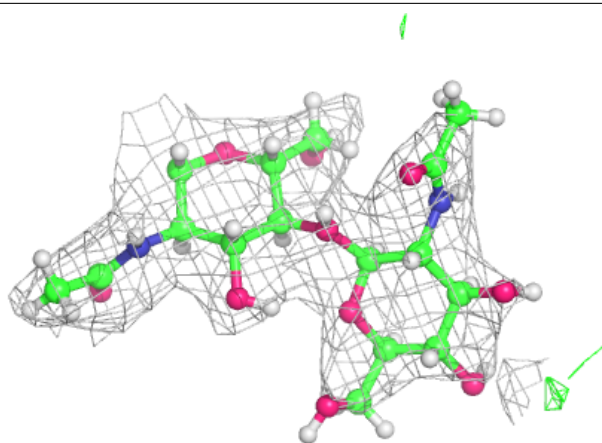
**Electron density around Chain F:**

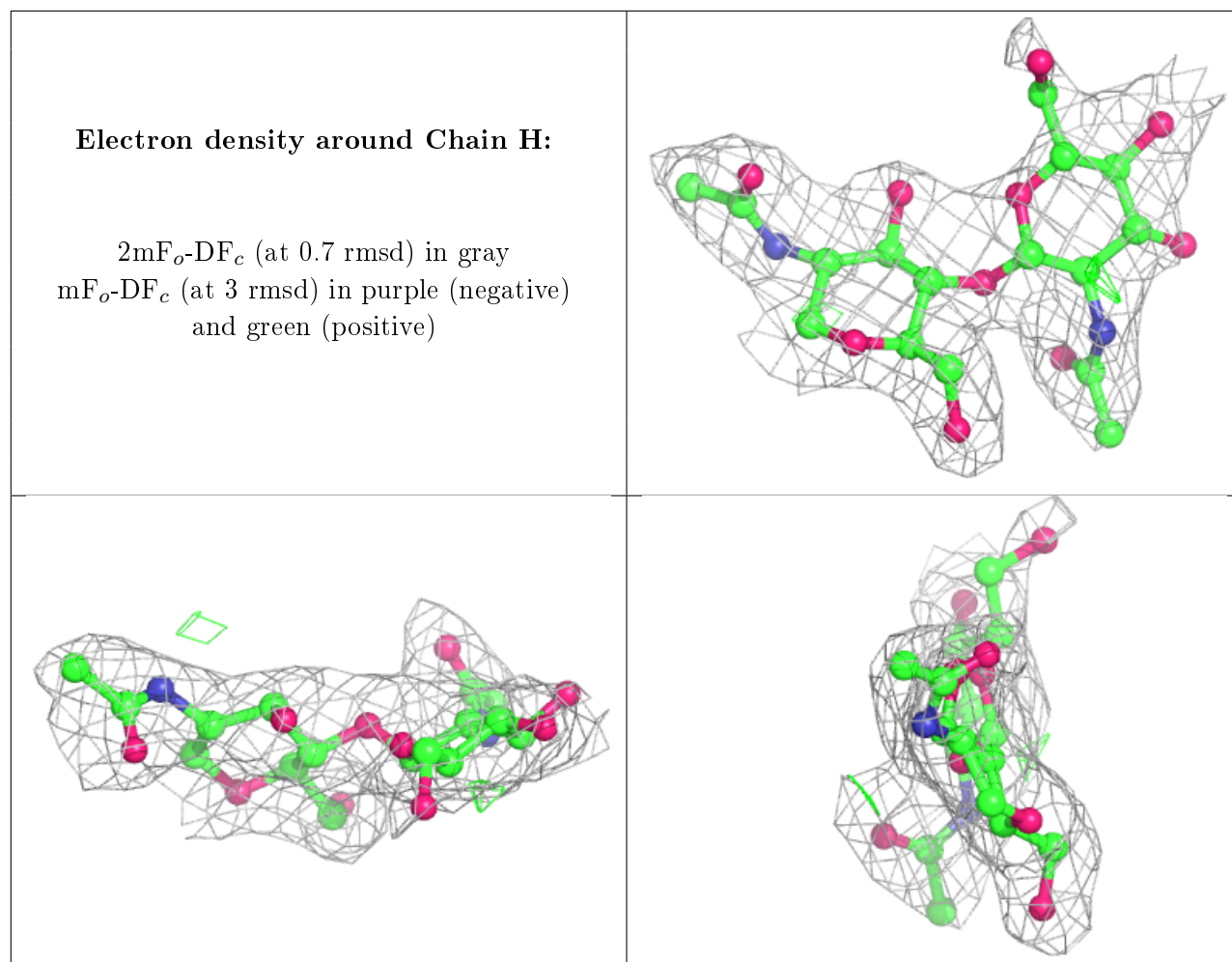
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**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
5	NAG	A	5010	14/15	0.77	0.28	70,85,102,104	0
5	NAG	A	5005	14/15	0.82	0.20	56,69,82,84	0
5	NAG	D	501	14/15	0.83	0.23	66,72,84,85	0
5	NAG	D	502	14/15	0.87	0.32	84,87,102,104	0
5	NAG	B	5009	14/15	0.87	0.19	52,64,70,72	0
5	NAG	B	5005	14/15	0.88	0.25	73,89,103,124	0
5	NAG	A	5007	14/15	0.88	0.16	51,62,74,75	0
5	NAG	C	2002	14/15	0.91	0.17	44,53,60,64	0
5	NAG	B	5002	14/15	0.91	0.18	50,65,81,82	0
5	NAG	C	2001	14/15	0.94	0.11	48,58,69,71	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	NAG	A	5006	14/15	0.94	0.12	37,41,48,51	0
5	NAG	B	5006	14/15	0.95	0.11	38,46,59,64	0
5	NAG	A	5002	14/15	0.96	0.13	29,39,48,53	0
4	ZN	B	5001	1/1	1.00	0.13	21,21,21,21	0
4	ZN	A	5001	1/1	1.00	0.15	21,21,21,21	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.