



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 18, 2023 – 11:53 PM EDT

PDB ID : 5BO6  
Title : Structure of human sialyltransferase ST8SiaIII in complex with CDP  
Authors : Volkers, G.; Worrall, L.; Strynadka, N.C.J.  
Deposited on : 2015-05-27  
Resolution : 2.07 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 : 2.35.1  
buster-report : 1.1.7 (2018)  
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

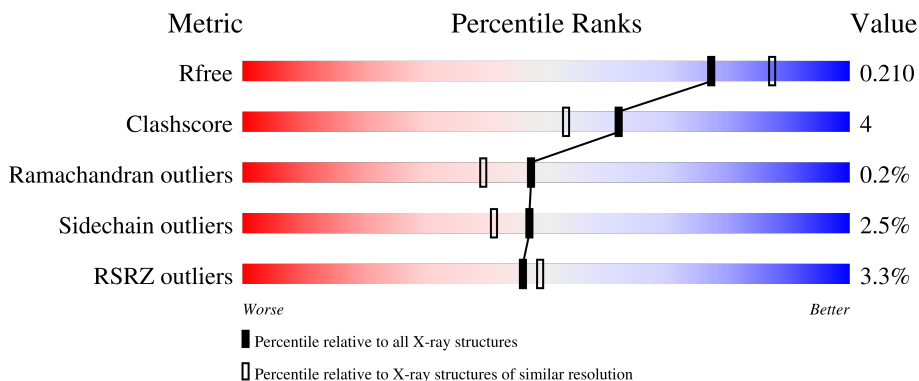
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.07 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.




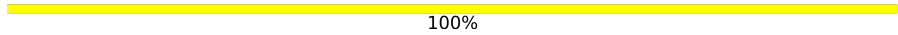

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2684 (2.08-2.04)
Clashscore	141614	2801 (2.08-2.04)
Ramachandran outliers	138981	2768 (2.08-2.04)
Sidechain outliers	138945	2768 (2.08-2.04)
RSRZ outliers	127900	2646 (2.08-2.04)

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	323	 2% 80% 8% • 10%
1	B	323	 3% 74% 10% • 14%
2	C	3	 33% 67%
3	D	2	 50% 50%
3	F	2	 50% 50%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	2	 50% 50%
3	H	2	 100%
4	E	3	 33% 67%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	NAG	B	406	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 5388 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 Sia-alpha-2,3-Gal-beta-1,4-GlcNAc-R:alpha 2,8-sialyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	291	2403	1553	418	422	10	0	0	0
1	B	277	2280	1473	397	400	10	0	0	0

There are 46 discrepancies between the modelled and reference sequences:

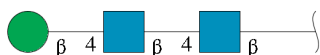
Chain	Residue	Modelled	Actual	Comment	Reference
A	58	ALA	-	expression tag	UNP O43173
A	59	PRO	-	expression tag	UNP O43173
A	60	GLU	-	expression tag	UNP O43173
A	61	HIS	-	expression tag	UNP O43173
A	62	HIS	-	expression tag	UNP O43173
A	63	HIS	-	expression tag	UNP O43173
A	64	HIS	-	expression tag	UNP O43173
A	65	HIS	-	expression tag	UNP O43173
A	66	HIS	-	expression tag	UNP O43173
A	67	ASP	-	expression tag	UNP O43173
A	68	TYR	-	expression tag	UNP O43173
A	69	ASP	-	expression tag	UNP O43173
A	70	ILE	-	expression tag	UNP O43173
A	71	PRO	-	expression tag	UNP O43173
A	72	THR	-	expression tag	UNP O43173
A	73	THR	-	expression tag	UNP O43173
A	74	GLU	-	expression tag	UNP O43173
A	75	ASN	-	expression tag	UNP O43173
A	76	LEU	-	expression tag	UNP O43173
A	77	TYR	-	expression tag	UNP O43173
A	78	PHE	-	expression tag	UNP O43173
A	79	GLN	-	expression tag	UNP O43173
A	80	GLY	-	expression tag	UNP O43173
B	58	ALA	-	expression tag	UNP O43173

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	59	PRO	-	expression tag	UNP O43173
B	60	GLU	-	expression tag	UNP O43173
B	61	HIS	-	expression tag	UNP O43173
B	62	HIS	-	expression tag	UNP O43173
B	63	HIS	-	expression tag	UNP O43173
B	64	HIS	-	expression tag	UNP O43173
B	65	HIS	-	expression tag	UNP O43173
B	66	HIS	-	expression tag	UNP O43173
B	67	ASP	-	expression tag	UNP O43173
B	68	TYR	-	expression tag	UNP O43173
B	69	ASP	-	expression tag	UNP O43173
B	70	ILE	-	expression tag	UNP O43173
B	71	PRO	-	expression tag	UNP O43173
B	72	THR	-	expression tag	UNP O43173
B	73	THR	-	expression tag	UNP O43173
B	74	GLU	-	expression tag	UNP O43173
B	75	ASN	-	expression tag	UNP O43173
B	76	LEU	-	expression tag	UNP O43173
B	77	TYR	-	expression tag	UNP O43173
B	78	PHE	-	expression tag	UNP O43173
B	79	GLN	-	expression tag	UNP O43173
B	80	GLY	-	expression tag	UNP O43173

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



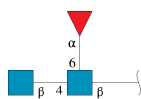
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	3	39	22	2	15	0	0	0

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



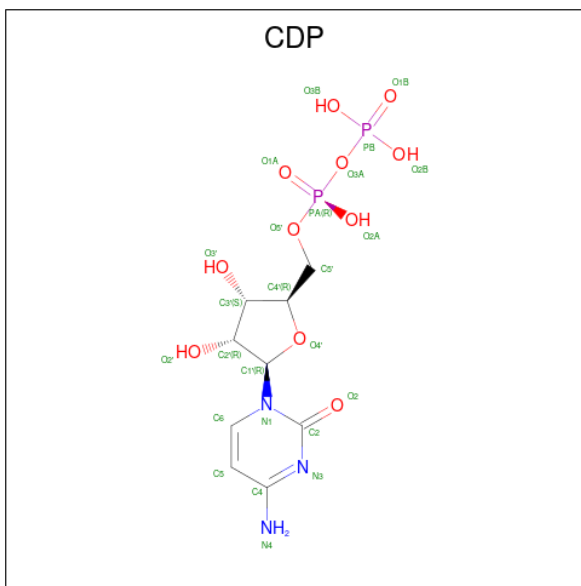
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	F	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	G	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	H	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	E	3	Total	C	N	O	0	0	0
			38	22	2	14			

- Molecule 5 is CYTIDINE-5'-DIPHOSPHATE (three-letter code: CDP) (formula: C<sub>9</sub>H<sub>15</sub>N<sub>3</sub>O<sub>11</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			25	9	3	11	2		
5	B	1	Total	C	N	O	P	0	0
			25	9	3	11	2		

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

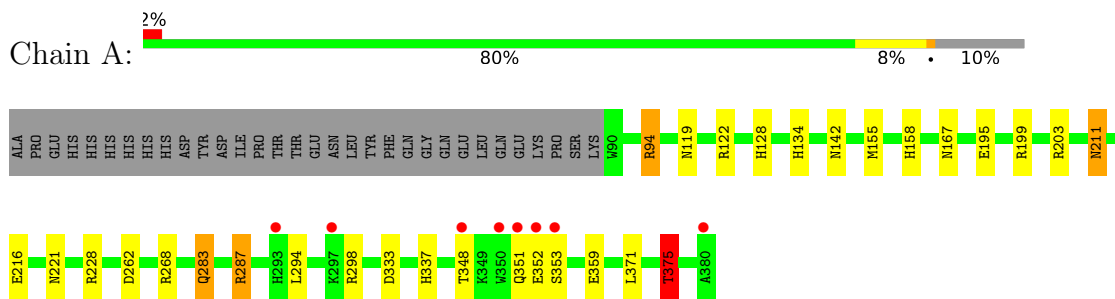
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	264	264	264	0	0
7	B	174	174	174	0	0

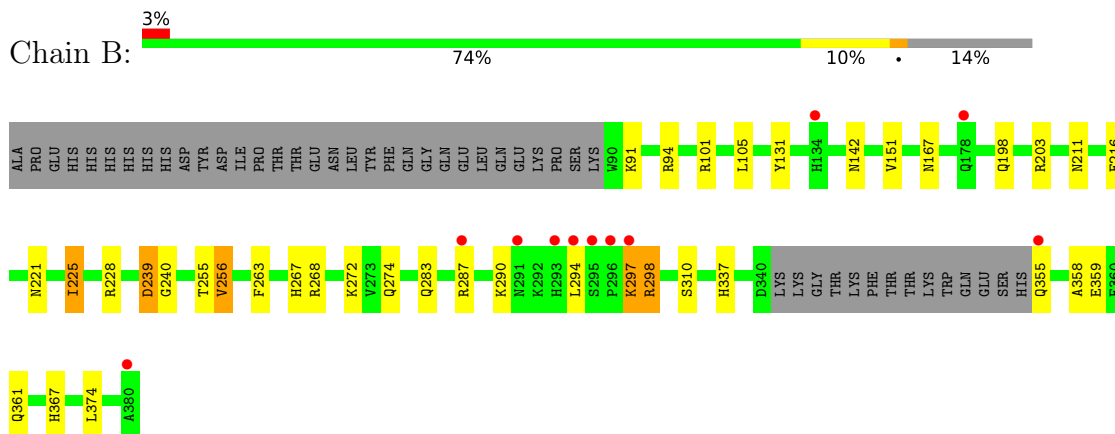
### 3 Residue-property plots [i](#)

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: Sia-alpha-2,3-Gal-beta-1,4-GlcNAc-R:alpha 2,8-sialyltransferase



- Molecule 1: Sia-alpha-2,3-Gal-beta-1,4-GlcNAc-R:alpha 2,8-sialyltransferase



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




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





MAG1  
MAG2

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

Chain F:  50% 50%

MAG1  
MAG2

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

Chain G:  50% 50%

MAG1  
MAG2

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

Chain H:  100%

MAG1  
MAG2

- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:  33% 67%

MAG1  
MAG2  
FUC3

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.93Å 96.87Å 124.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.26 – 2.07 47.45 – 2.07	Depositor EDS
% Data completeness (in resolution range)	98.4 (58.26-2.07) 98.4 (47.45-2.07)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.37 (at 2.07Å)	Xtrriage
Refinement program	REFMAC 5.8.0123	Depositor
R, $R_{free}$	0.168 , 0.202 0.177 , 0.210	Depositor DCC
$R_{free}$ test set	2670 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.8	Xtrriage
Anisotropy	0.487	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 47.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5388	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.22% 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: BMA, NAG, CDP, FUC

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	1.12	4/2474 (0.2%)	1.01	11/3352 (0.3%)
1	B	1.03	1/2345 (0.0%)	1.03	11/3177 (0.3%)
All	All	1.08	5/4819 (0.1%)	1.02	22/6529 (0.3%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	375	THR	CB-CG2	-6.35	1.31	1.52
1	B	142	ASN	CB-CG	6.33	1.65	1.51
1	A	94	ARG	CZ-NH1	6.22	1.41	1.33
1	A	211	ASN	CB-CG	6.12	1.65	1.51
1	A	142	ASN	CB-CG	5.75	1.64	1.51

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	268	ARG	NE-CZ-NH1	10.58	125.59	120.30
1	B	268	ARG	NE-CZ-NH2	-10.55	115.03	120.30
1	A	268	ARG	NE-CZ-NH2	-8.99	115.80	120.30
1	B	268	ARG	NE-CZ-NH1	8.10	124.35	120.30
1	A	94	ARG	NE-CZ-NH2	-6.93	116.84	120.30
1	B	272	LYS	CD-CE-NZ	6.63	126.94	111.70
1	A	287	ARG	NE-CZ-NH1	6.59	123.59	120.30
1	B	94	ARG	NE-CZ-NH2	-6.40	117.10	120.30
1	B	203	ARG	NE-CZ-NH1	6.38	123.49	120.30
1	A	122	ARG	NE-CZ-NH1	6.01	123.31	120.30
1	B	94	ARG	NE-CZ-NH1	5.97	123.29	120.30
1	B	228	ARG	NE-CZ-NH1	5.97	123.29	120.30
1	B	225	ILE	CG1-CB-CG2	5.90	124.39	111.40
1	B	256	VAL	CB-CA-C	-5.76	100.46	111.40

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	228	ARG	NE-CZ-NH2	-5.67	117.47	120.30
1	A	94	ARG	CD-NE-CZ	5.58	131.41	123.60
1	A	333	ASP	CB-CG-OD1	5.24	123.02	118.30
1	A	262	ASP	CB-CG-OD1	5.20	122.98	118.30
1	A	268	ARG	CD-NE-CZ	5.19	130.86	123.60
1	A	203	ARG	N-CA-CB	-5.09	101.44	110.60
1	B	268	ARG	CG-CD-NE	-5.05	101.19	111.80
1	B	198	GLN	CB-CA-C	-5.01	100.38	110.40

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	2403	0	2341	18	0
1	B	2280	0	2223	23	0
2	C	39	0	34	0	0
3	D	28	0	25	0	0
3	F	28	0	25	2	0
3	G	28	0	25	0	0
3	H	28	0	25	0	0
4	E	38	0	34	0	0
5	A	25	0	12	0	0
5	B	25	0	12	0	0
6	B	28	0	26	0	0
7	A	264	0	0	14	1
7	B	174	0	0	7	1
All	All	5388	0	4782	43	1

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.

All (43) 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:155:MET:HE2	7:A:763:HOH:O	1.70	0.92
1:A:119:ASN:HB2	7:A:657:HOH:O	1.75	0.87
1:B:355:GLN:HG3	1:B:358:ALA:HB3	1.59	0.82
1:B:367:HIS:HD2	1:B:374:LEU:H	1.28	0.79
1:B:239:ASP:CG	1:B:240:GLY:H	1.90	0.74
1:A:199:ARG:HD3	7:A:671:HOH:O	1.89	0.73
1:A:94:ARG:NH2	7:A:501:HOH:O	2.20	0.70
1:B:287:ARG:NH1	7:B:502:HOH:O	2.27	0.68
1:B:255:THR:HG22	7:B:553:HOH:O	1.95	0.66
1:A:375:THR:HG23	7:A:637:HOH:O	1.95	0.66
1:A:158:HIS:CE1	7:A:535:HOH:O	2.51	0.64
7:A:658:HOH:O	3:F:1:NAG:H83	1.98	0.64
1:B:355:GLN:HG3	1:B:358:ALA:CB	2.28	0.64
1:A:211:ASN:HB2	7:A:712:HOH:O	1.98	0.62
1:A:94:ARG:NE	7:A:501:HOH:O	2.06	0.61
1:B:297:LYS:HG3	1:B:298:ARG:HB2	1.82	0.61
1:B:101:ARG:HD3	1:B:310:SER:O	2.00	0.61
1:B:367:HIS:CD2	1:B:374:LEU:H	2.14	0.59
1:B:283:GLN:HG3	7:B:638:HOH:O	2.02	0.58
1:A:216:GLU:HG3	1:A:221:ASN:HD21	1.67	0.58
1:B:239:ASP:CG	1:B:240:GLY:N	2.52	0.57
1:A:375:THR:CG2	7:A:637:HOH:O	2.54	0.56
1:B:105:LEU:HD22	7:B:634:HOH:O	2.07	0.53
1:B:274:GLN:HG3	7:B:630:HOH:O	2.08	0.52
1:B:297:LYS:HG3	1:B:298:ARG:HD3	1.93	0.50
7:A:502:HOH:O	3:F:1:NAG:O6	2.20	0.48
1:B:216:GLU:HG3	1:B:221:ASN:HD21	1.78	0.48
1:A:294:LEU:HD21	1:A:359:GLU:HG2	1.95	0.47
1:A:211:ASN:ND2	7:A:510:HOH:O	2.48	0.46
1:B:225:ILE:HD12	1:B:225:ILE:H	1.81	0.46
1:A:283:GLN:HG2	7:A:506:HOH:O	2.15	0.45
1:A:348:THR:OG1	1:A:351:GLN:NE2	2.50	0.44
1:A:351:GLN:O	1:A:353:SER:N	2.49	0.44
1:B:263:PHE:O	1:B:267:HIS:HD2	2.00	0.44
1:B:105:LEU:HG	7:B:601:HOH:O	2.17	0.43
1:B:239:ASP:OD2	1:B:240:GLY:N	2.28	0.43
1:A:128:HIS:CE1	1:A:134:HIS:CE1	3.07	0.42
1:B:294:LEU:HD21	1:B:359:GLU:HG2	2.02	0.42
1:A:167:ASN:O	1:A:337:HIS:HA	2.20	0.42
1:B:167:ASN:O	1:B:337:HIS:HA	2.20	0.41
1:B:211:ASN:OD1	7:B:501:HOH:O	2.21	0.41
1:A:195:GLU:HB2	7:A:693:HOH:O	2.20	0.40

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:131:TYR:CE1	1:B:298:ARG:HG3	2.56	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:561:HOH:O	7:B:565:HOH:O[4_455]	2.11	0.09

## 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	289/323 (90%)	284 (98%)	4 (1%)	1 (0%)	41	32
1	B	273/323 (84%)	269 (98%)	4 (2%)	0	100	100
All	All	562/646 (87%)	553 (98%)	8 (1%)	1 (0%)	47	39

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	352	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	263/294 (90%)	258 (98%)	5 (2%)	57	53
1	B	250/294 (85%)	242 (97%)	8 (3%)	39	32
All	All	513/588 (87%)	500 (98%)	13 (2%)	47	41

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	283	GLN
1	A	287	ARG
1	A	298	ARG
1	A	371	LEU
1	A	375	THR
1	B	91	LYS
1	B	151	VAL
1	B	239	ASP
1	B	256	VAL
1	B	290	LYS
1	B	297	LYS
1	B	298	ARG
1	B	361	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	119	ASN
1	A	128	HIS
1	A	134	HIS
1	A	156	ASN
1	A	221	ASN
1	A	351	GLN
1	A	378	HIS
1	B	221	ASN
1	B	229	ASN
1	B	251	HIS
1	B	267	HIS
1	B	274	GLN
1	B	367	HIS
1	B	378	HIS

### 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](#)

14 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	NAG	C	1	2,1	14,14,15	0.51	0	17,19,21	1.31	2 (11%)
2	NAG	C	2	2	14,14,15	0.63	0	17,19,21	1.06	1 (5%)
2	BMA	C	3	2	11,11,12	0.50	0	15,15,17	0.86	0
3	NAG	D	1	3,1	14,14,15	0.62	0	17,19,21	0.93	0
3	NAG	D	2	3	14,14,15	0.42	0	17,19,21	1.27	3 (17%)
4	NAG	E	1	4,1	14,14,15	0.82	1 (7%)	17,19,21	0.93	1 (5%)
4	NAG	E	2	4	14,14,15	0.68	0	17,19,21	1.37	3 (17%)
4	FUC	E	3	4	10,10,11	0.57	0	14,14,16	0.90	0
3	NAG	F	1	3,1	14,14,15	1.05	1 (7%)	17,19,21	1.81	5 (29%)
3	NAG	F	2	3	14,14,15	0.64	0	17,19,21	1.39	1 (5%)
3	NAG	G	1	3,1	14,14,15	0.52	0	17,19,21	1.17	1 (5%)
3	NAG	G	2	3	14,14,15	0.54	0	17,19,21	0.94	0
3	NAG	H	1	3,1	14,14,15	0.78	0	17,19,21	1.66	4 (23%)
3	NAG	H	2	3	14,14,15	0.67	0	17,19,21	0.98	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
2	NAG	C	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	C	2	2	-	2/6/23/26	0/1/1/1
2	BMA	C	3	2	-	0/2/19/22	0/1/1/1
3	NAG	D	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	D	2	3	-	1/6/23/26	0/1/1/1
4	NAG	E	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	E	2	4	-	0/6/23/26	0/1/1/1
4	FUC	E	3	4	-	-	0/1/1/1
3	NAG	F	1	3,1	-	2/6/23/26	0/1/1/1
3	NAG	F	2	3	-	0/6/23/26	0/1/1/1
3	NAG	G	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	G	2	3	-	0/6/23/26	0/1/1/1
3	NAG	H	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	H	2	3	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	1	NAG	C8-C7	-3.13	1.44	1.50
4	E	1	NAG	C1-C2	2.53	1.56	1.52

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	1	NAG	C1-O5-C5	4.20	117.89	112.19
3	F	1	NAG	O5-C1-C2	-3.71	105.44	111.29
3	F	1	NAG	C2-N2-C7	3.31	127.61	122.90
3	F	1	NAG	C1-O5-C5	3.29	116.66	112.19
2	C	1	NAG	C1-C2-N2	3.13	115.84	110.49
3	D	2	NAG	O5-C1-C2	-2.88	106.74	111.29
4	E	2	NAG	O5-C5-C6	2.74	111.50	107.20
3	D	2	NAG	C1-C2-N2	2.71	115.11	110.49
3	H	1	NAG	C1-C2-N2	-2.63	106.00	110.49
4	E	1	NAG	C1-C2-N2	2.55	114.84	110.49
2	C	2	NAG	O5-C1-C2	-2.53	107.30	111.29
3	H	1	NAG	O4-C4-C3	-2.48	104.62	110.35
3	G	1	NAG	O5-C5-C6	2.44	111.03	107.20
3	F	2	NAG	C2-N2-C7	2.41	126.34	122.90
3	F	1	NAG	O6-C6-C5	-2.33	103.30	111.29
3	H	1	NAG	O5-C5-C6	2.17	110.61	107.20
3	H	2	NAG	O4-C4-C5	2.16	114.66	109.30
2	C	1	NAG	O5-C5-C6	2.09	110.48	107.20
4	E	2	NAG	C1-C2-N2	-2.08	106.93	110.49

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	1	NAG	O3-C3-C2	-2.07	105.17	109.47
4	E	2	NAG	O5-C5-C4	-2.06	105.81	110.83
3	D	2	NAG	C1-O5-C5	2.02	114.94	112.19

There are no chirality outliers.

All (5) torsion outliers are listed below:

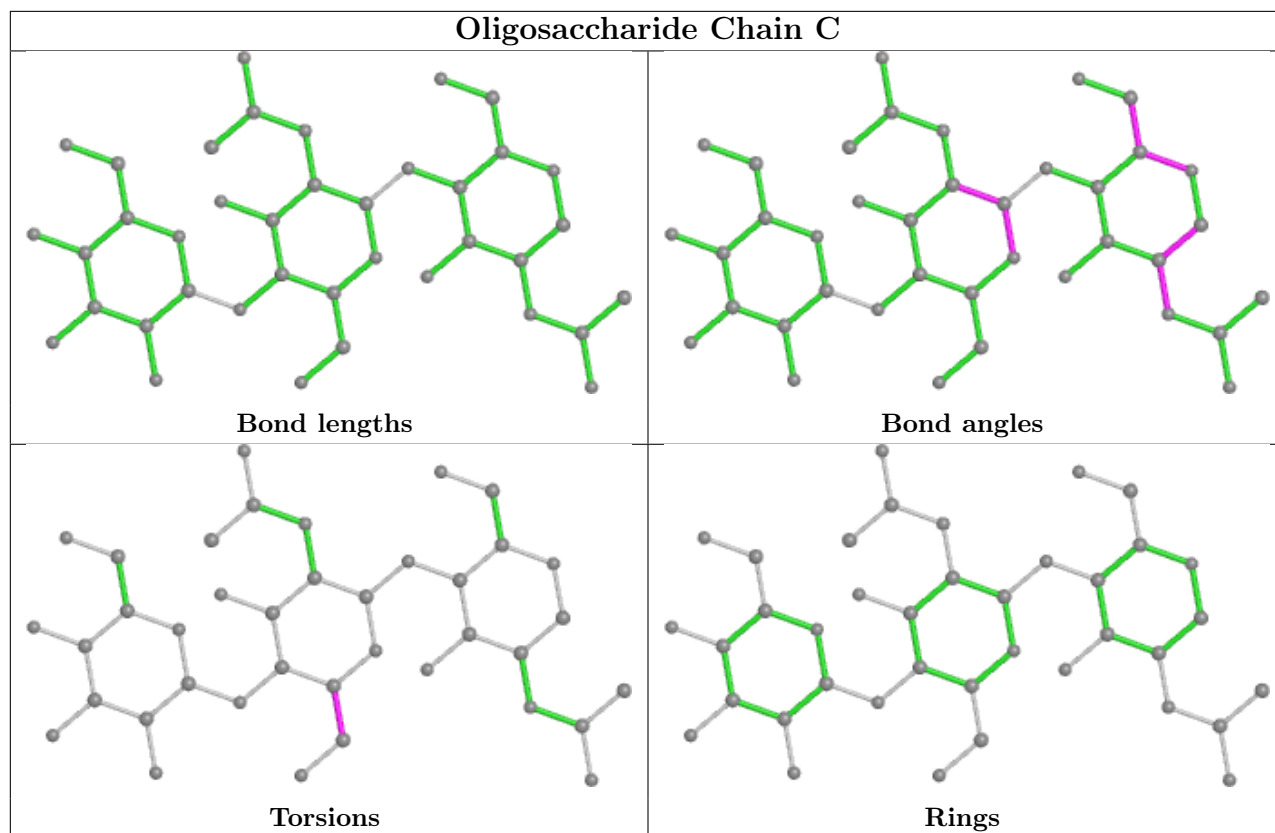
Mol	Chain	Res	Type	Atoms
2	C	2	NAG	O5-C5-C6-O6
2	C	2	NAG	C4-C5-C6-O6
3	F	1	NAG	C8-C7-N2-C2
3	F	1	NAG	O7-C7-N2-C2
3	D	2	NAG	O5-C5-C6-O6

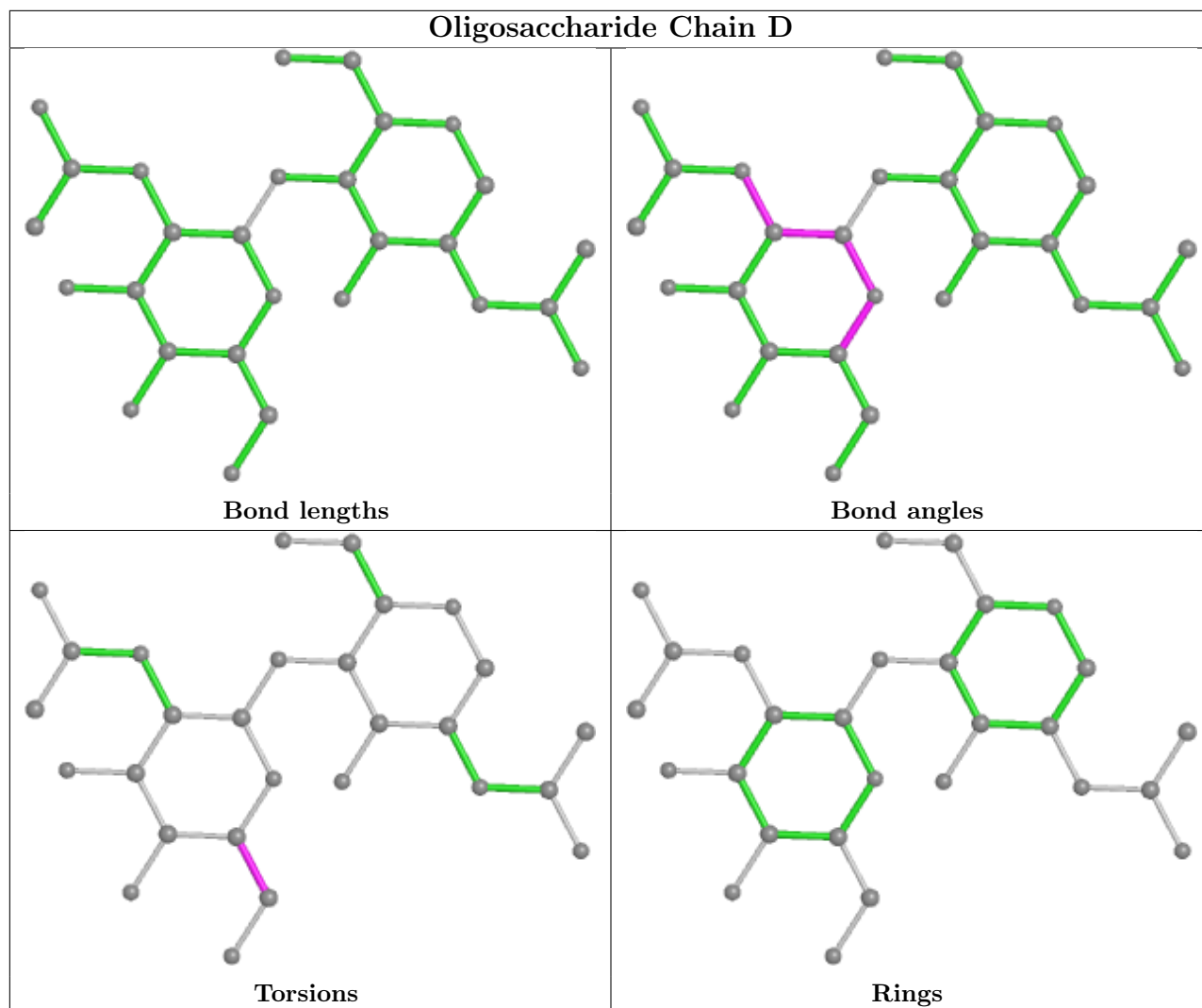
There are no ring outliers.

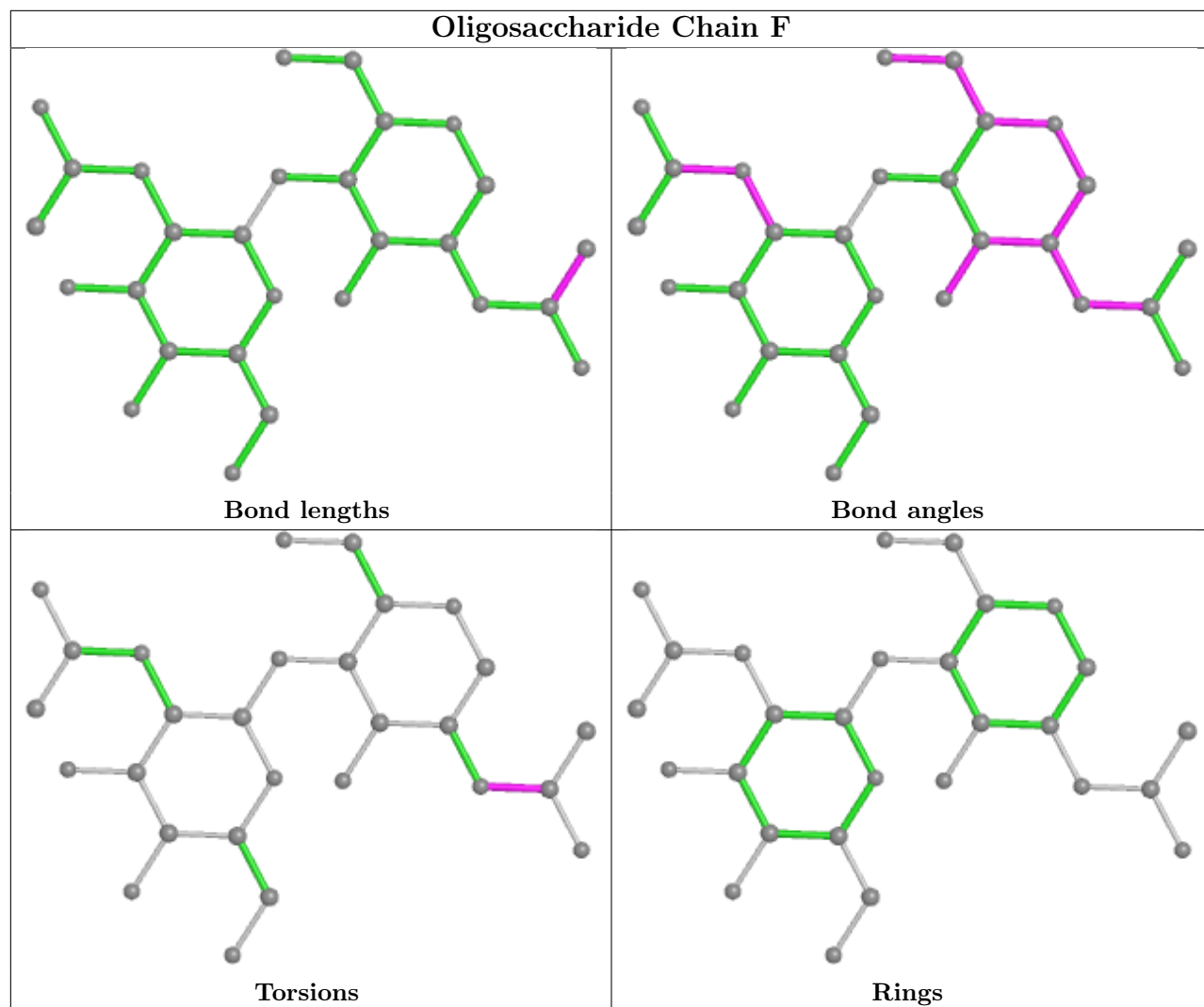
1 monomer is involved in 2 short contacts:

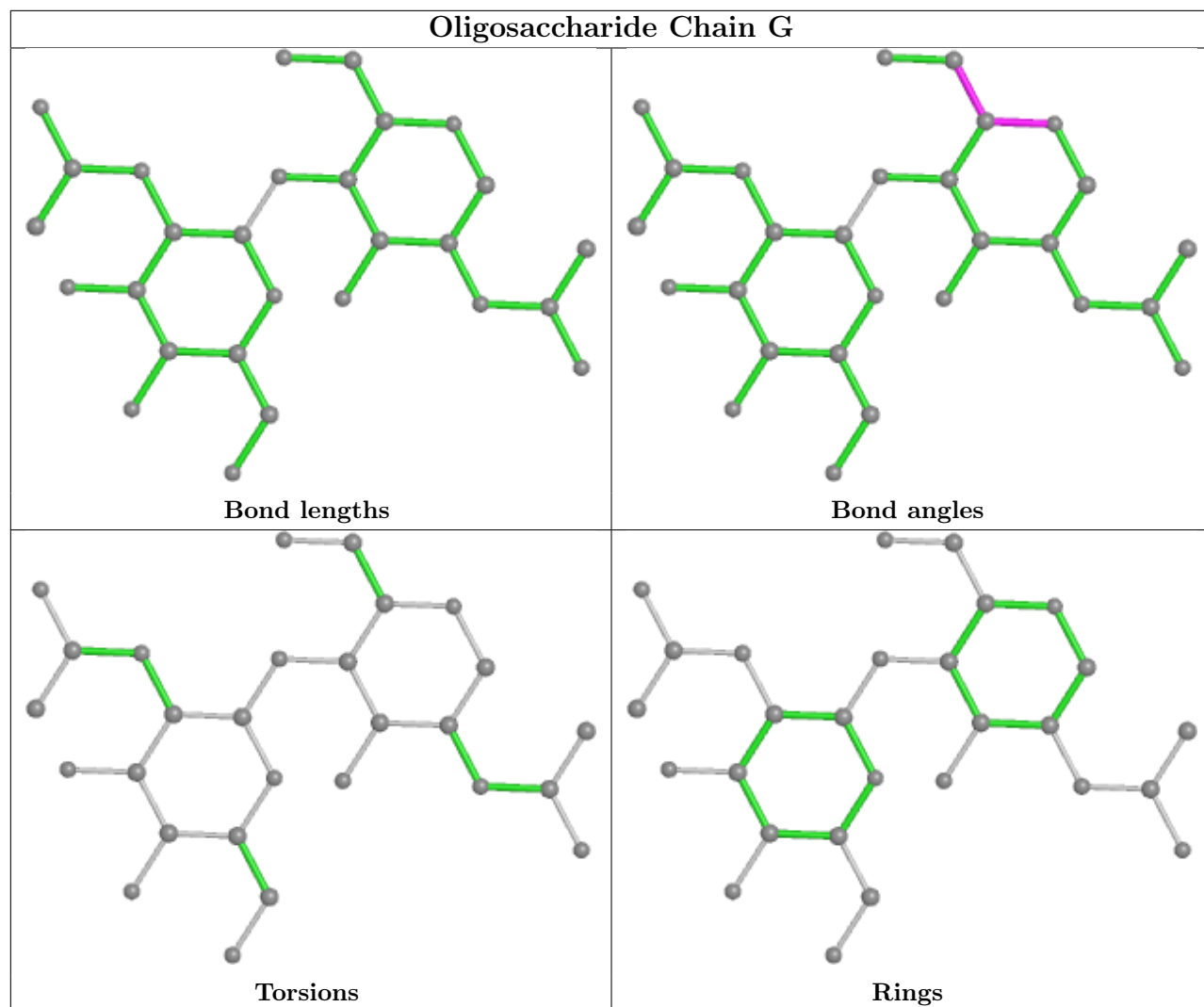
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	1	NAG	2	0

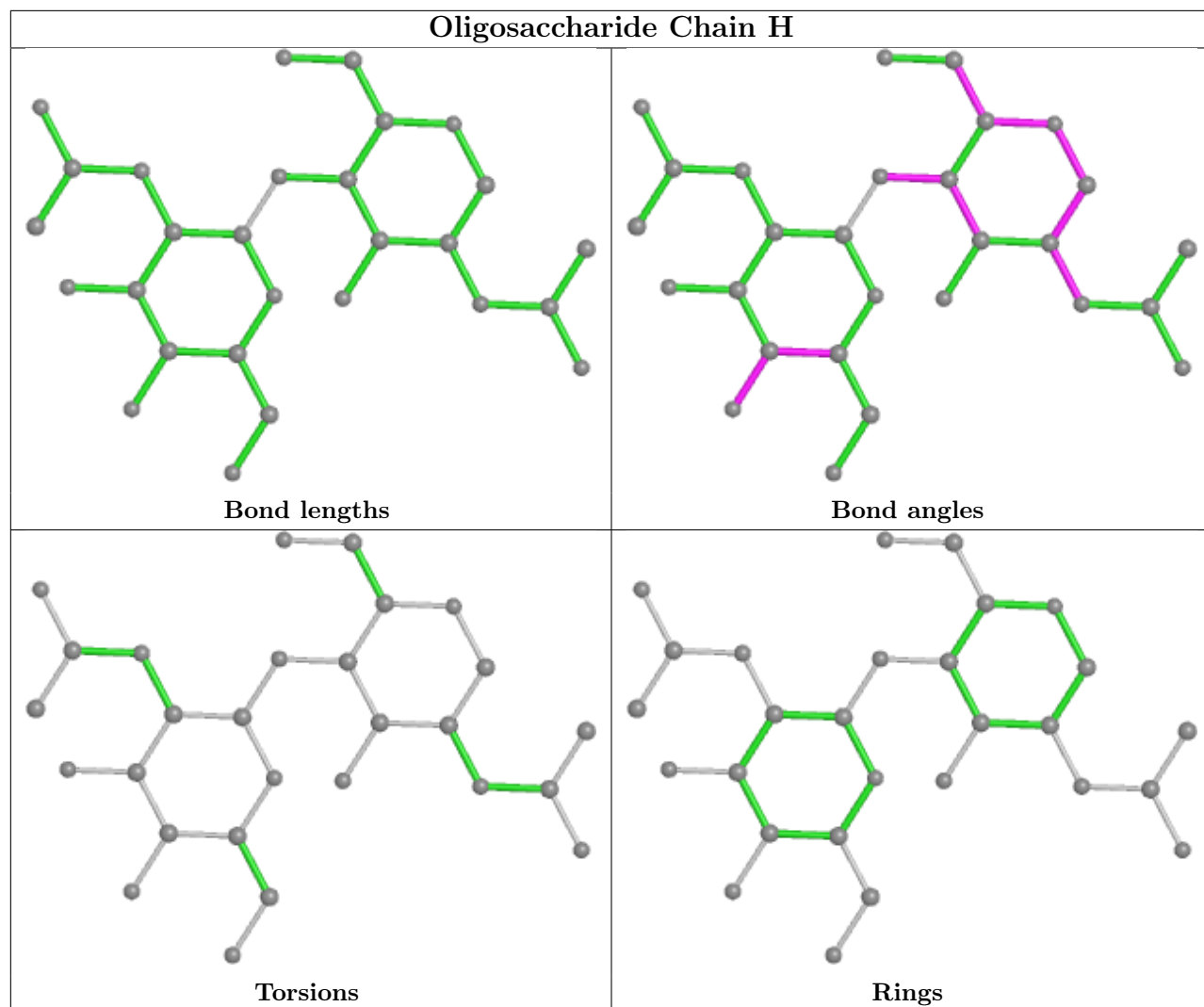
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

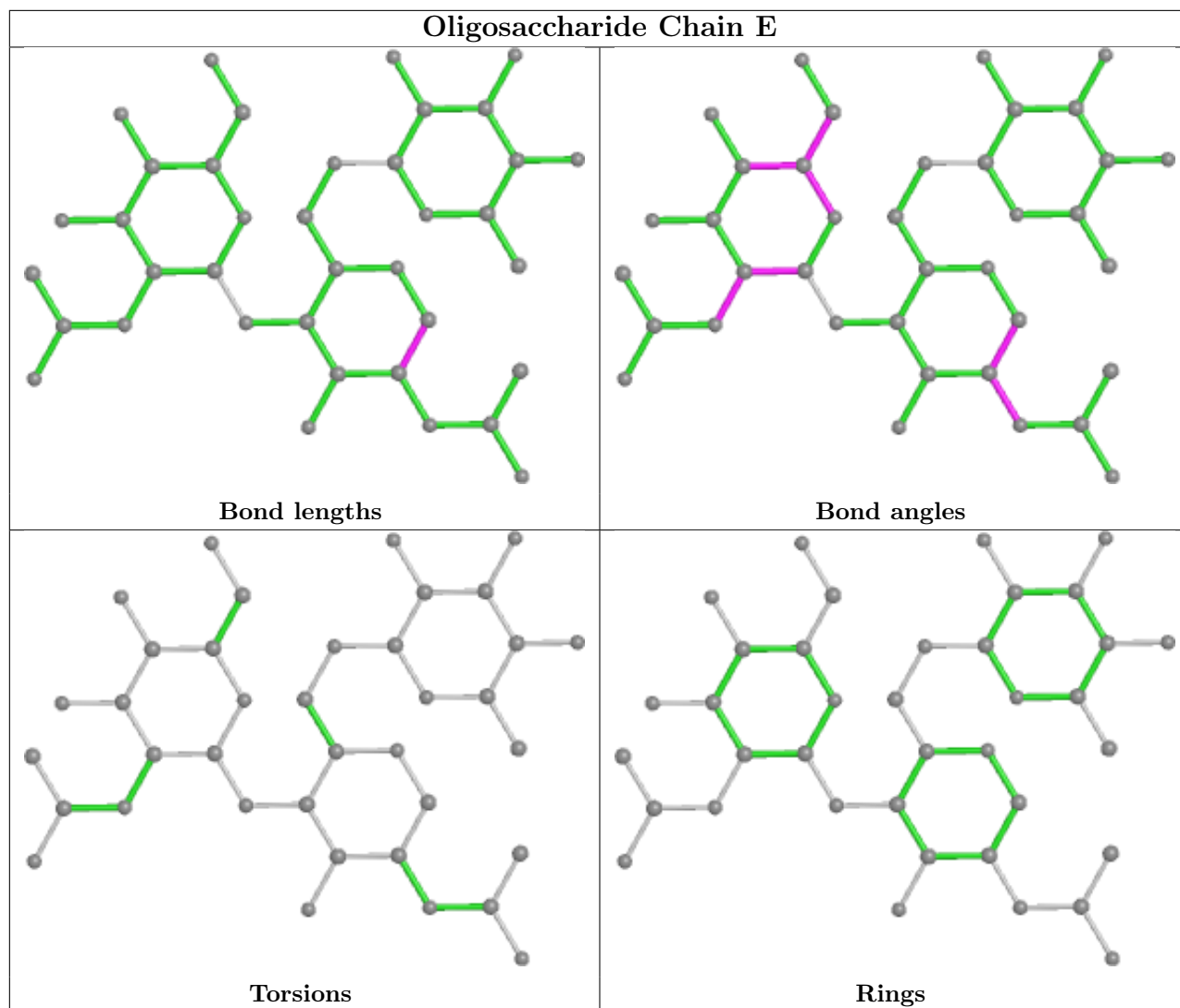












## 5.6 Ligand geometry [i](#)

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
5	CDP	A	411	-	24,26,26	0.78	1 (4%)	37,40,40	1.16	5 (13%)
6	NAG	B	406	1	14,14,15	1.16	1 (7%)	17,19,21	1.44	3 (17%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	CDP	B	407	-	24,26,26	0.81	1 (4%)	37,40,40	1.12	3 (8%)
6	NAG	B	405	1	14,14,15	0.57	0	17,19,21	2.49	5 (29%)

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	CDP	A	411	-	-	2/16/32/32	0/2/2/2
6	NAG	B	406	1	-	1/6/23/26	0/1/1/1
5	CDP	B	407	-	-	6/16/32/32	0/2/2/2
6	NAG	B	405	1	-	2/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	406	NAG	C1-C2	3.14	1.57	1.52
5	B	407	CDP	C6-C5	2.19	1.40	1.35
5	A	411	CDP	C6-C5	2.15	1.40	1.35

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	405	NAG	C1-O5-C5	5.78	120.02	112.19
6	B	405	NAG	O5-C1-C2	-5.13	103.18	111.29
6	B	405	NAG	C8-C7-N2	3.87	122.66	116.10
6	B	406	NAG	C1-C2-N2	3.11	115.81	110.49
5	A	411	CDP	O3B-PB-O1B	3.00	122.42	110.68
6	B	406	NAG	O5-C5-C6	2.90	111.75	107.20
5	A	411	CDP	O2B-PB-O3B	2.45	116.99	107.64
5	B	407	CDP	O2-C2-N3	-2.39	118.45	122.33
6	B	406	NAG	C2-N2-C7	2.31	126.20	122.90
5	A	411	CDP	O3A-PB-O1B	-2.31	98.36	111.19
5	B	407	CDP	O2A-PA-O1A	2.31	123.65	112.24
6	B	405	NAG	O7-C7-N2	-2.24	117.83	121.95
6	B	405	NAG	C3-C4-C5	-2.18	106.35	110.24
5	A	411	CDP	C2'-C1'-N1	2.13	119.24	113.22
5	B	407	CDP	O3'-C3'-C4'	2.07	117.03	111.05
5	A	411	CDP	O4'-C1'-N1	-2.04	103.71	108.36

There are no chirality outliers.

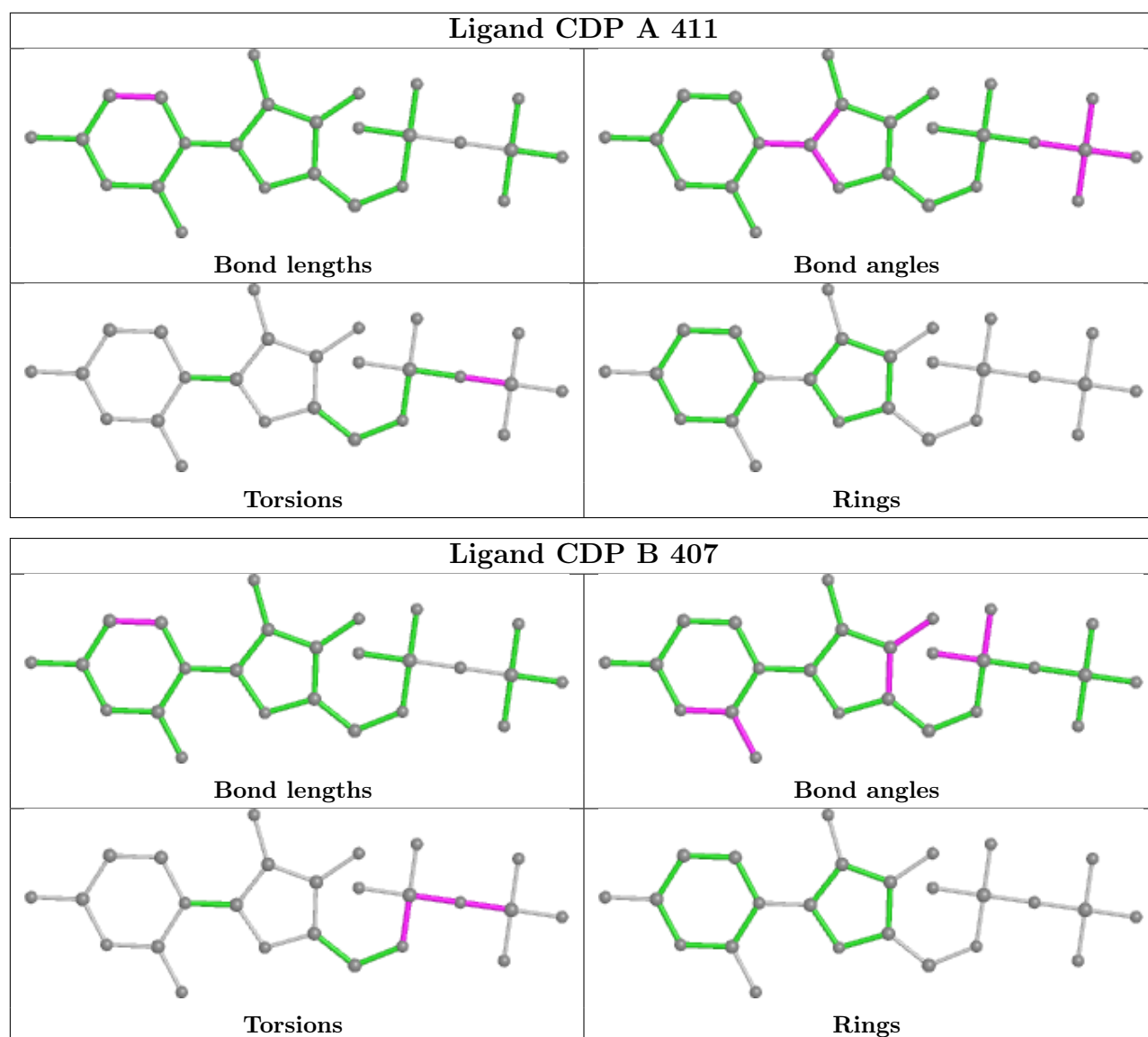
All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	411	CDP	PA-O3A-PB-O3B
5	B	407	CDP	C5'-O5'-PA-O1A
6	B	405	NAG	C8-C7-N2-C2
6	B	405	NAG	O7-C7-N2-C2
5	A	411	CDP	PA-O3A-PB-O1B
6	B	406	NAG	C1-C2-N2-C7
5	B	407	CDP	PA-O3A-PB-O1B
5	B	407	CDP	PA-O3A-PB-O3B
5	B	407	CDP	PA-O3A-PB-O2B
5	B	407	CDP	C5'-O5'-PA-O3A
5	B	407	CDP	PB-O3A-PA-O1A

There are no ring outliers.

No monomer is involved in short contacts.

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



## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [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	291/323 (90%)	-0.11	8 (2%) 54 57	21, 32, 54, 91	0
1	B	277/323 (85%)	0.09	11 (3%) 38 40	24, 40, 66, 100	0
All	All	568/646 (87%)	-0.01	19 (3%) 46 49	21, 36, 61, 100	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	350	TRP	4.8
1	B	380	ALA	4.8
1	B	355	GLN	4.0
1	A	352	GLU	3.9
1	A	353	SER	3.8
1	B	291	ASN	3.6
1	B	293	HIS	3.5
1	A	351	GLN	3.5
1	B	297	LYS	3.4
1	A	348	THR	3.3
1	B	287	ARG	3.0
1	B	295	SER	2.5
1	B	294	LEU	2.4
1	A	297	LYS	2.4
1	A	293	HIS	2.3
1	B	296	PRO	2.2
1	B	134	HIS	2.2
1	B	178	GLN	2.1
1	A	380	ALA	2.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

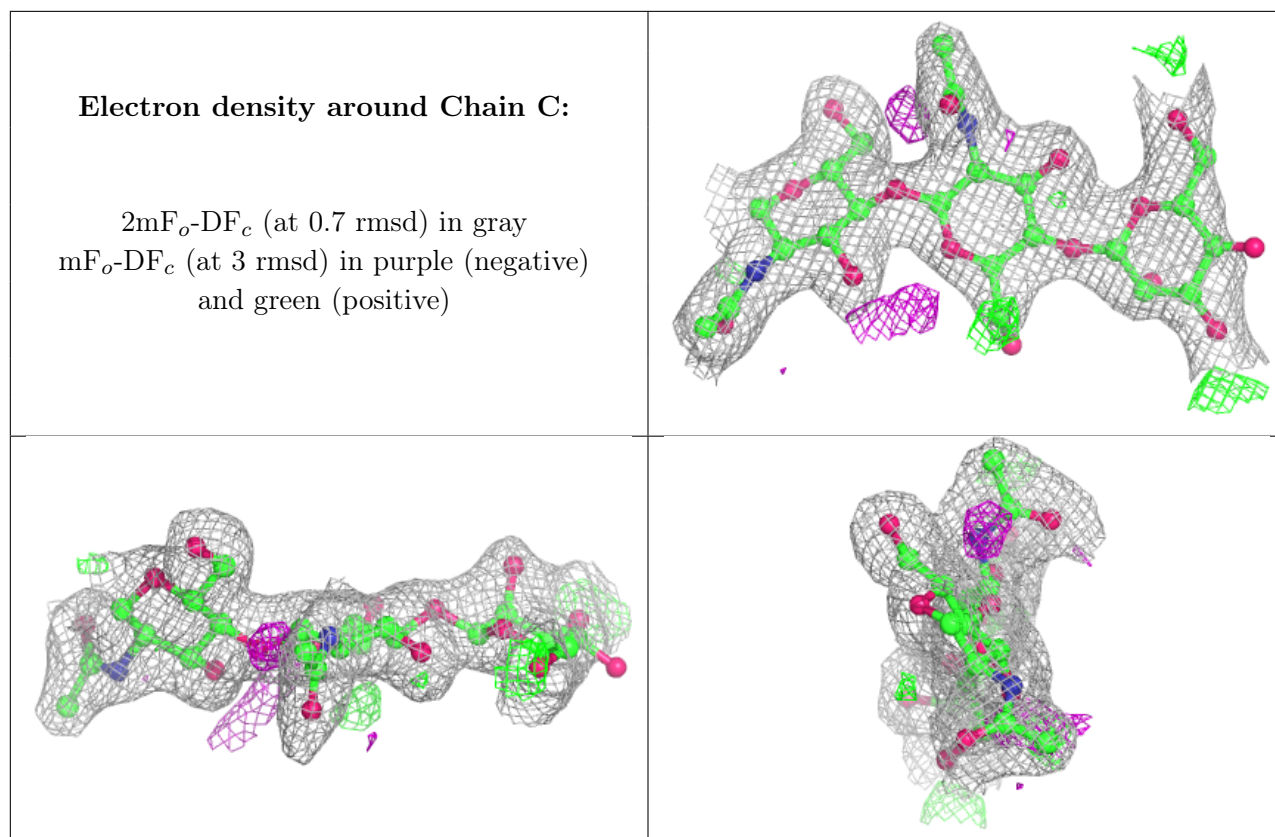
There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

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

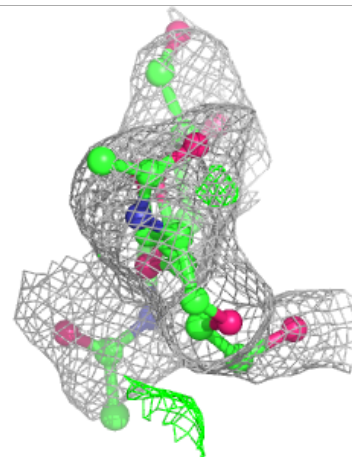
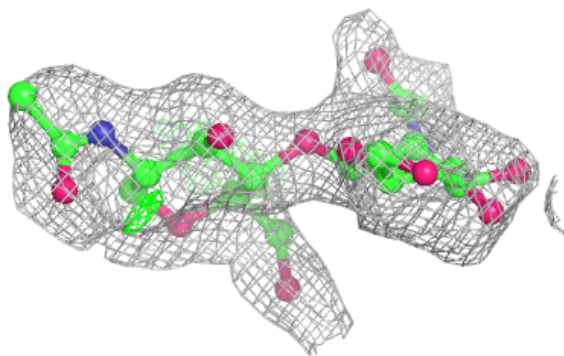
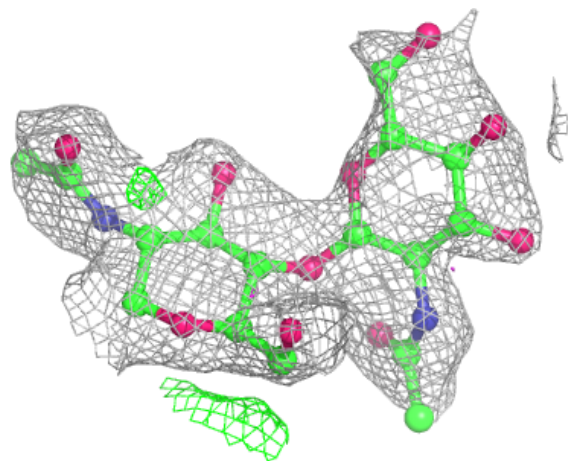
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	BMA	C	3	11/12	0.74	0.20	69,76,81,86	0
3	NAG	D	1	14/15	0.83	0.29	57,71,78,87	0
3	NAG	D	2	14/15	0.84	0.39	72,92,98,99	0
3	NAG	H	2	14/15	0.89	0.34	57,63,73,77	0
3	NAG	G	2	14/15	0.90	0.17	44,51,65,69	0
2	NAG	C	2	14/15	0.92	0.12	40,47,67,72	0
3	NAG	F	2	14/15	0.92	0.14	44,51,64,70	0
4	NAG	E	2	14/15	0.93	0.15	47,52,64,65	0
4	NAG	E	1	14/15	0.94	0.10	27,31,36,40	0
3	NAG	F	1	14/15	0.94	0.12	25,30,39,40	0
3	NAG	H	1	14/15	0.95	0.08	34,41,44,56	0
3	NAG	G	1	14/15	0.95	0.09	28,33,37,44	0
4	FUC	E	3	10/11	0.95	0.10	38,42,45,47	0
2	NAG	C	1	14/15	0.96	0.09	31,35,41,47	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.



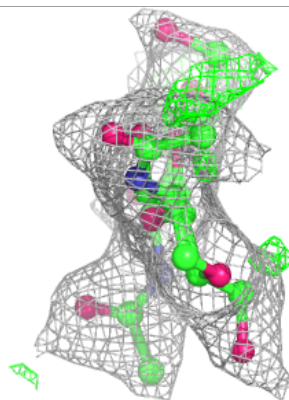
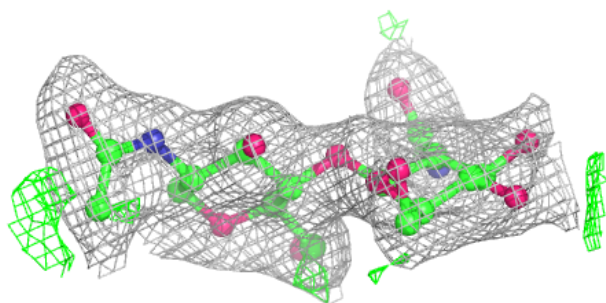
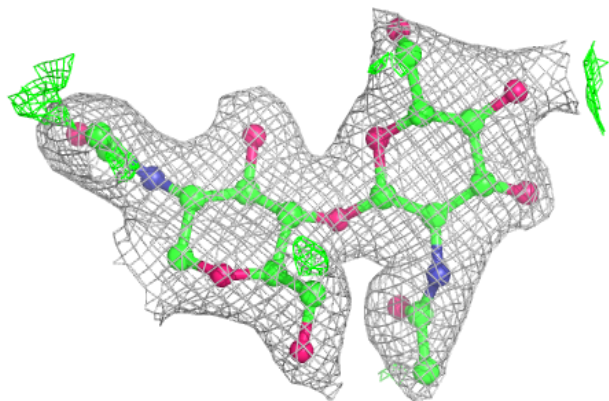
**Electron density around Chain D:**

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



**Electron density around Chain F:**

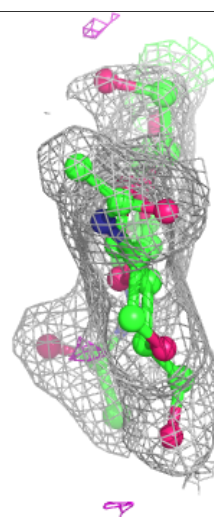
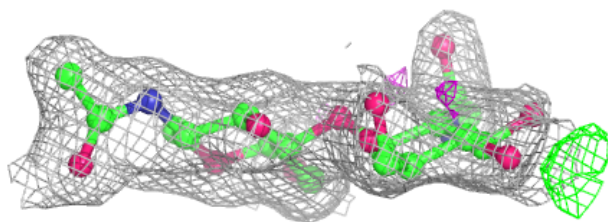
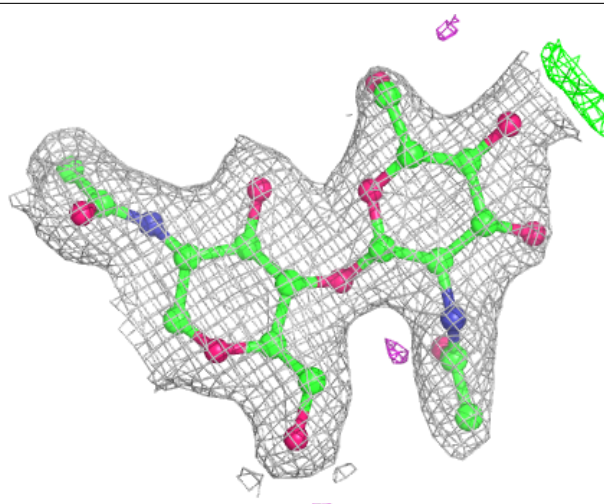
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

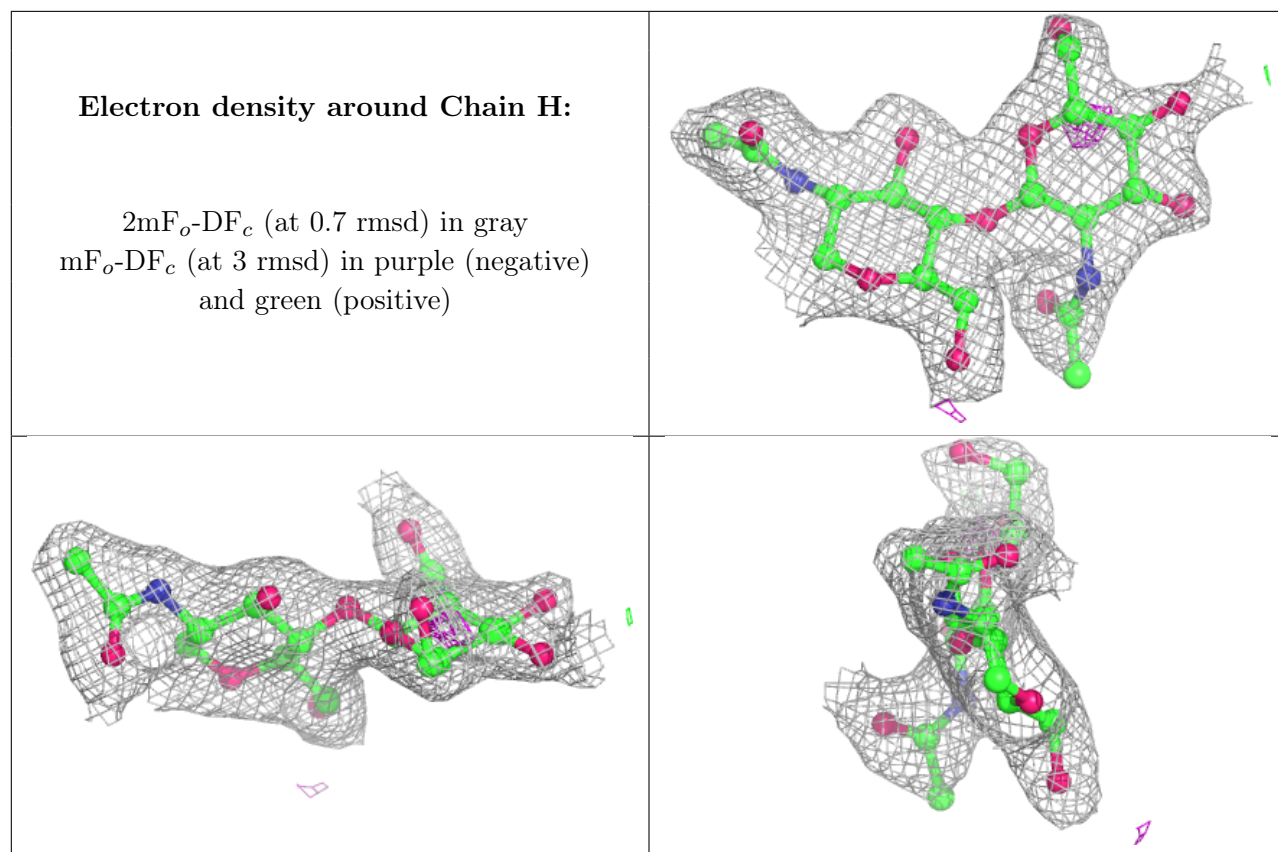


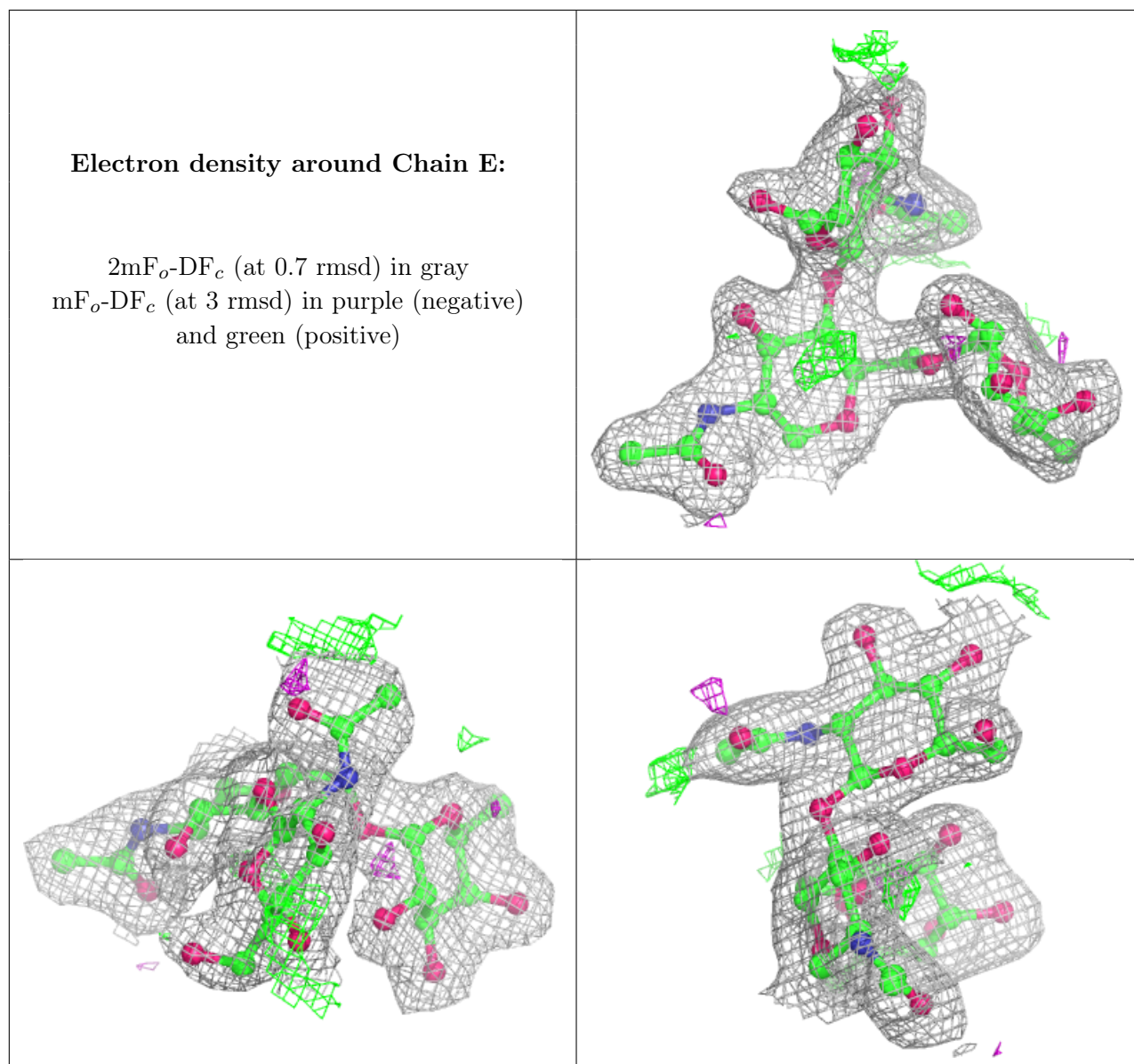


**Electron density around Chain G:**

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





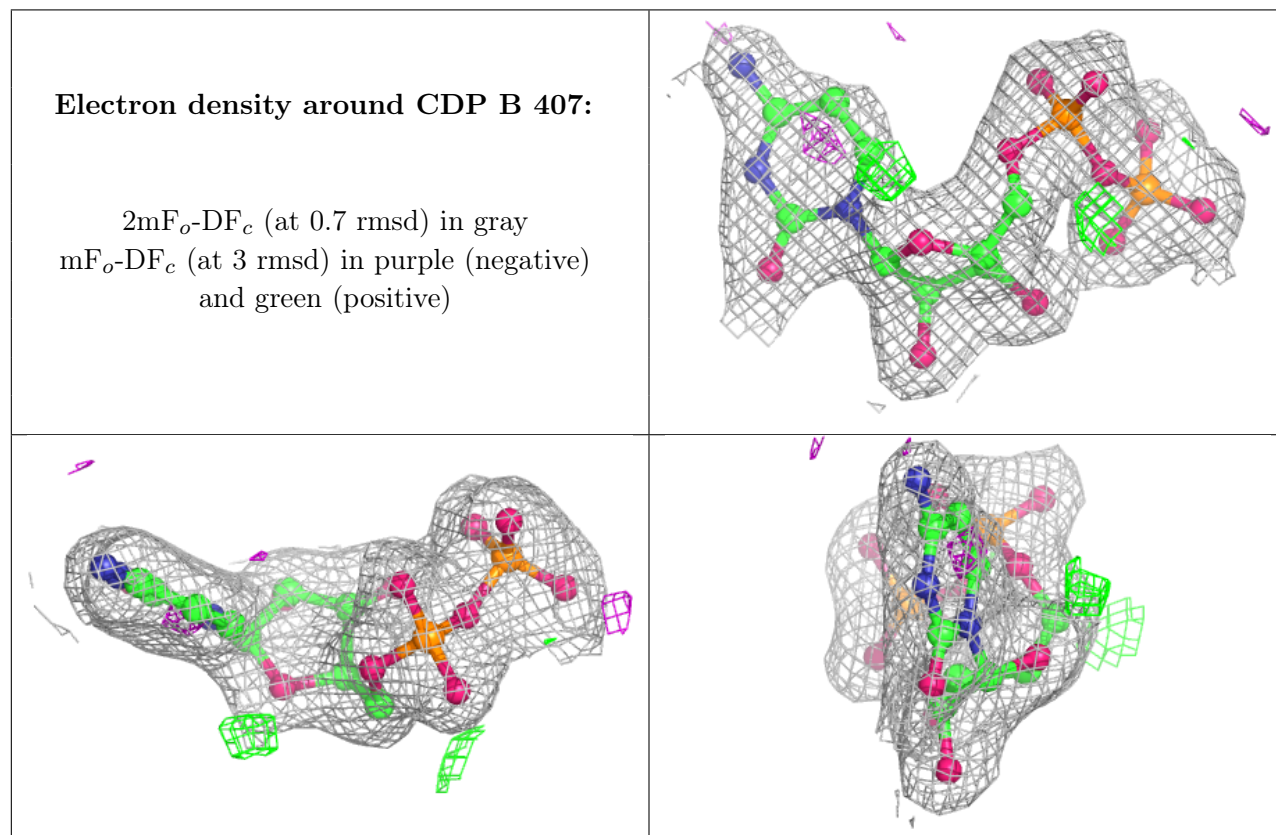


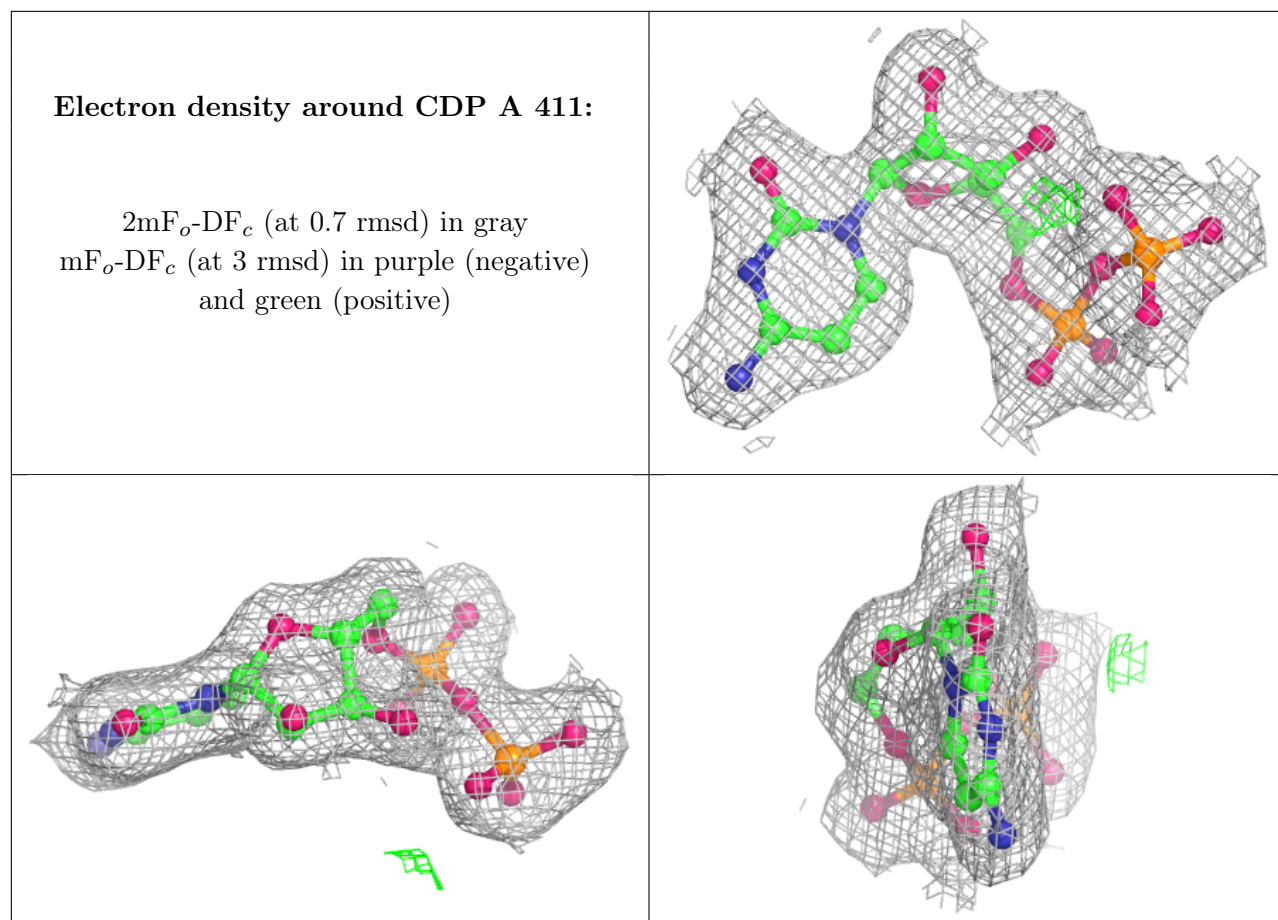
## 6.4 Ligands [i](#)

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	B	406	14/15	0.69	0.40	71,89,100,105	0
6	NAG	B	405	14/15	0.79	0.29	58,68,76,79	0
5	CDP	B	407	25/25	0.96	0.09	37,44,50,52	0
5	CDP	A	411	25/25	0.98	0.09	25,27,35,38	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.





## 6.5 Other polymers [i](#)

There are no such residues in this entry.