



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 6, 2023 – 09:31 AM EDT

PDB ID : 4EEO  
Title : Crystal structure of human M340H-beta-1,4-galactosyltransferase-1 (M340H-B4GAL-T1) in complex with GLCNAC-BETA1,6-GlcNAc-ALPHA-benzyl  
l  
Authors : Ramakrishnan, B.; Qasba, P.K.  
Deposited on : 2012-03-28  
Resolution : 2.30 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  
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

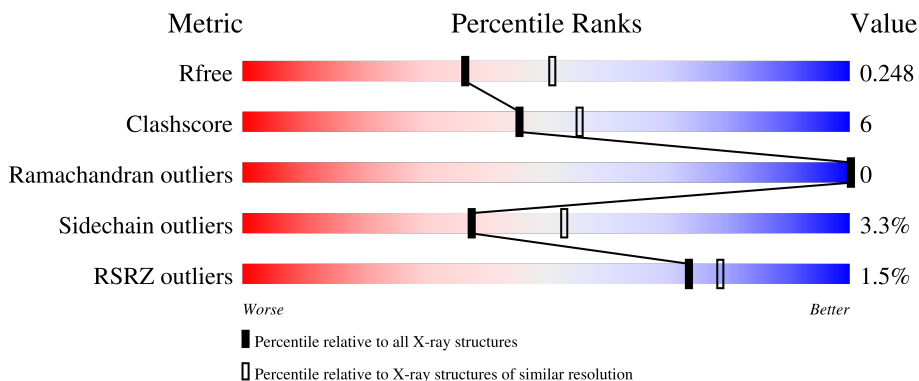
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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.




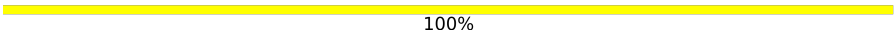
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	287	<div style="display: flex; align-items: center;"> <div style="width: 2%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 82%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 12%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 20px;">82%      12%      • 5%</p>
1	B	287	<div style="display: flex; align-items: center;"> <div style="width: 80%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 13%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 20px;">80%      13%      • 5%</p>
1	C	287	<div style="display: flex; align-items: center;"> <div style="width: 2%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 78%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 16%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="margin-left: 20px;">2%      78%      16%      • 5%</p>
2	D	2	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 10px; background-color: yellow;"></div> </div> <p style="margin-left: 20px;">100%</p>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	E	2	 50% 50%
2	F	2	 100%

## 2 Entry composition i

There are 7 unique types of molecules in this entry. The entry contains 7408 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 Beta-1,4-galactosyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	273	2234	1430	388	405	11	0	3	0
1	B	273	2234	1431	386	405	12	0	3	0
1	C	273	2218	1420	386	401	11	0	0	0

There are 51 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	112	ALA	-	expression tag	UNP P15291
A	113	SER	-	expression tag	UNP P15291
A	114	MET	-	expression tag	UNP P15291
A	115	THR	-	expression tag	UNP P15291
A	116	GLY	-	expression tag	UNP P15291
A	117	GLY	-	expression tag	UNP P15291
A	118	GLN	-	expression tag	UNP P15291
A	119	GLN	-	expression tag	UNP P15291
A	120	MET	-	expression tag	UNP P15291
A	121	GLY	-	expression tag	UNP P15291
A	122	ARG	-	expression tag	UNP P15291
A	123	GLY	-	expression tag	UNP P15291
A	124	SER	-	expression tag	UNP P15291
A	125	ALA	-	expression tag	UNP P15291
A	337	THR	ARG	engineered mutation	UNP P15291
A	338	THR	CYS	engineered mutation	UNP P15291
A	340	HIS	MET	engineered mutation	UNP P15291
B	112	ALA	-	expression tag	UNP P15291
B	113	SER	-	expression tag	UNP P15291
B	114	MET	-	expression tag	UNP P15291
B	115	THR	-	expression tag	UNP P15291
B	116	GLY	-	expression tag	UNP P15291
B	117	GLY	-	expression tag	UNP P15291

*Continued on next page...*

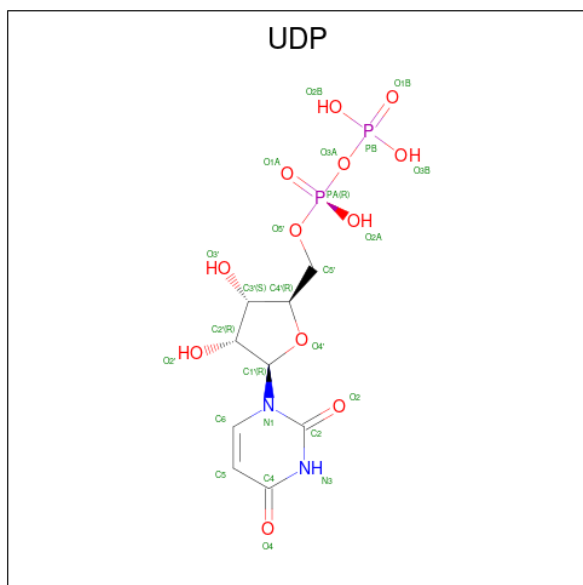
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	118	GLN	-	expression tag	UNP P15291
B	119	GLN	-	expression tag	UNP P15291
B	120	MET	-	expression tag	UNP P15291
B	121	GLY	-	expression tag	UNP P15291
B	122	ARG	-	expression tag	UNP P15291
B	123	GLY	-	expression tag	UNP P15291
B	124	SER	-	expression tag	UNP P15291
B	125	ALA	-	expression tag	UNP P15291
B	337	THR	ARG	engineered mutation	UNP P15291
B	338	THR	CYS	engineered mutation	UNP P15291
B	340	HIS	MET	engineered mutation	UNP P15291
C	112	ALA	-	expression tag	UNP P15291
C	113	SER	-	expression tag	UNP P15291
C	114	MET	-	expression tag	UNP P15291
C	115	THR	-	expression tag	UNP P15291
C	116	GLY	-	expression tag	UNP P15291
C	117	GLY	-	expression tag	UNP P15291
C	118	GLN	-	expression tag	UNP P15291
C	119	GLN	-	expression tag	UNP P15291
C	120	MET	-	expression tag	UNP P15291
C	121	GLY	-	expression tag	UNP P15291
C	122	ARG	-	expression tag	UNP P15291
C	123	GLY	-	expression tag	UNP P15291
C	124	SER	-	expression tag	UNP P15291
C	125	ALA	-	expression tag	UNP P15291
C	337	THR	ARG	engineered mutation	UNP P15291
C	338	THR	CYS	engineered mutation	UNP P15291
C	340	HIS	MET	engineered mutation	UNP P15291

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-6)-benzyl 2-acetamido-2-deoxy-alpha-D-glucopyranoside.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	2	Total	C	N	O	0	0	0
			36	23	2	11			
2	E	2	Total	C	N	O	0	0	0
			36	23	2	11			
2	F	2	Total	C	N	O	0	0	0
			36	23	2	11			

- Molecule 3 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: C<sub>9</sub>H<sub>14</sub>N<sub>2</sub>O<sub>12</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
3	B	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
3	C	1	Total	C	N	O	P	0	0
			25	9	2	12	2		

- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

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

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	C	1	Total O S 5 4 1	0	0
5	C	1	Total O S 5 4 1	0	0

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



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

- Molecule 7 is water.

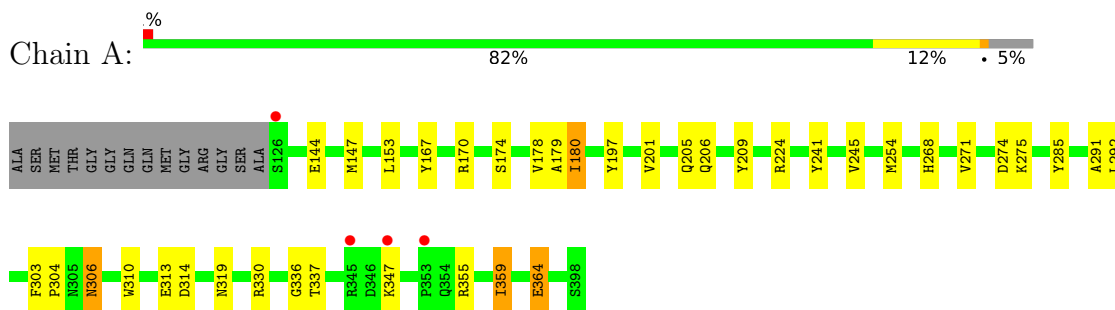
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	171	Total O 171 171	0	0
7	B	197	Total O 197 197	0	0
7	C	95	Total O 95 95	0	0



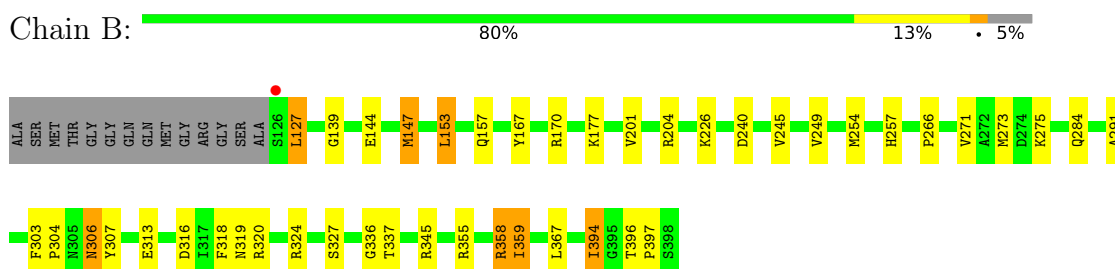
### 3 Residue-property plots [i](#)

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.

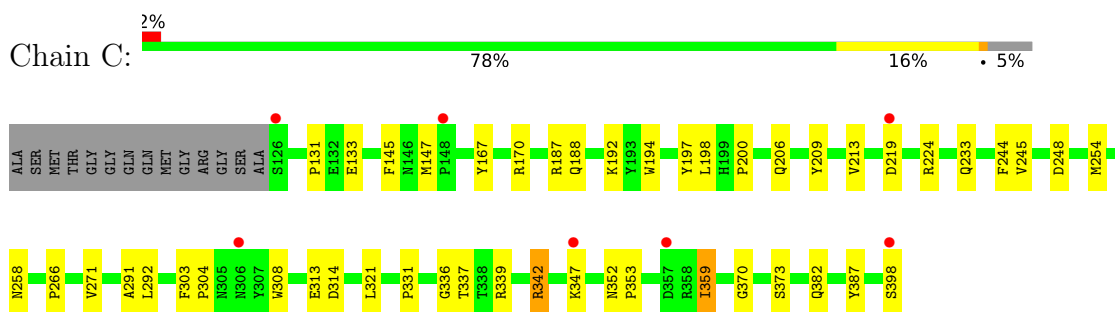
- Molecule 1: Beta-1,4-galactosyltransferase 1



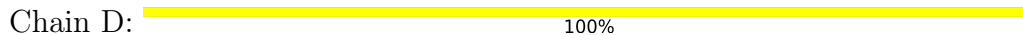
- Molecule 1: Beta-1,4-galactosyltransferase 1



- Molecule 1: Beta-1,4-galactosyltransferase 1



- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-6)-benzyl 2-acetamido-2-deoxy-alpha-D-glucopyranoside



BBV1  
MAG2

- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-6)-benzyl 2-acetamido-2-deoxy-alpha-D-glucopyranoside

Chain E:  50% 50%BBV1  
MAG2

- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-6)-benzyl 2-acetamido-2-deoxy-alpha-D-glucopyranoside

Chain F:  100%BBV1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.80Å 198.22Å 143.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.80 – 2.30 40.80 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.5 (40.80-2.30) 99.5 (40.80-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.30 (at 2.29Å)	Xtrriage
Refinement program	REFMAC 5.5.0110	Depositor
R, $R_{free}$	0.203 , 0.255 0.200 , 0.248	Depositor DCC
$R_{free}$ test set	3420 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.1	Xtrriage
Anisotropy	0.138	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 34.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.000 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.009 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7408	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.19% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: UDP, BBV, MN, NAG, GOL, SO4

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	0.87	1/2305 (0.0%)	0.81	1/3131 (0.0%)
1	B	0.96	1/2305 (0.0%)	0.86	3/3131 (0.1%)
1	C	0.82	0/2280	0.77	0/3097
All	All	0.88	2/6890 (0.0%)	0.81	4/9359 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	201	VAL	CB-CG2	5.74	1.65	1.52
1	A	364	GLU	CG-CD	5.59	1.60	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	358	ARG	NE-CZ-NH1	-5.74	117.43	120.30
1	B	204	ARG	NE-CZ-NH1	-5.45	117.58	120.30
1	B	127	LEU	CA-CB-CG	5.33	127.55	115.30
1	A	180	ILE	CG1-CB-CG2	-5.01	100.38	111.40

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	2234	0	2180	26	0
1	B	2234	0	2181	30	0
1	C	2218	0	2162	30	0
2	D	36	0	33	0	0
2	E	36	0	33	1	0
2	F	36	0	33	0	0
3	A	25	0	11	0	0
3	B	25	0	11	1	0
3	C	25	0	11	2	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
5	A	20	0	0	0	0
5	B	25	0	0	0	0
5	C	10	0	0	1	0
6	A	6	0	8	2	0
6	B	6	0	8	2	0
6	C	6	0	8	0	0
7	A	171	0	0	0	0
7	B	197	0	0	1	0
7	C	95	0	0	2	0
All	All	7408	0	6679	87	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 6.

All (87) 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)
6:B:410:GOL:H31	7:B:695:HOH:O	1.67	0.94
1:B:139:GLY:HA2	1:B:257:HIS:HD2	1.49	0.75
1:A:306:ASN:H	1:A:306:ASN:HD22	1.34	0.73
1:B:139:GLY:HA2	1:B:257:HIS:CD2	2.33	0.64
1:B:303:PHE:HB3	1:B:304:PRO:HD2	1.80	0.62
1:A:310:TRP:CD1	6:A:409:GOL:H12	2.35	0.62
1:A:319:ASN:HD21	1:A:359:ILE:HD12	1.65	0.61
1:A:144:GLU:O	1:A:197:TYR:OH	2.17	0.59
1:B:284:GLN:HG2	1:B:318:PHE:CZ	2.38	0.59
1:C:254:MET:HG2	1:C:337:THR:HG22	1.83	0.59
1:C:133:GLU:HG3	7:C:577:HOH:O	2.03	0.58
1:A:306:ASN:HD22	1:A:306:ASN:N	2.01	0.57
1:B:306:ASN:H	1:B:306:ASN:HD22	1.53	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:131:PRO:HD2	1:C:206:GLN:HE22	1.71	0.56
1:B:144[A]:GLU:OE1	1:B:147[A]:MET:HE3	2.06	0.55
1:A:291:ALA:C	1:A:292:LEU:HD12	2.27	0.55
1:A:254:MET:SD	1:A:337:THR:HG22	2.47	0.54
1:C:188:GLN:O	1:C:192:LYS:HG3	2.08	0.54
1:B:271:VAL:HG22	1:B:336:GLY:HA3	1.90	0.54
1:C:308:TRP:O	1:C:352:ASN:HB2	2.08	0.54
1:B:324:ARG:NH1	1:B:367:LEU:HD21	2.22	0.53
1:A:292:LEU:HD12	1:A:292:LEU:N	2.23	0.53
1:A:303:PHE:HB3	1:A:304:PRO:HD2	1.90	0.53
1:A:271:VAL:HG22	1:A:336:GLY:HA3	1.91	0.53
1:B:144[A]:GLU:OE1	1:B:147[A]:MET:CE	2.57	0.52
1:A:254:MET:SD	1:A:337:THR:CG2	2.97	0.52
1:C:271:VAL:HG22	1:C:336:GLY:HA3	1.92	0.52
1:B:266:PRO:HD2	1:B:327:SER:O	2.09	0.52
1:C:248:ASP:HB3	3:C:403:UDP:O3'	2.10	0.51
1:B:319:ASN:HD21	1:B:359:ILE:HD12	1.75	0.51
1:A:254:MET:HB2	1:A:337:THR:HG22	1.93	0.51
1:A:355:ARG:O	1:A:359:ILE:HG23	2.11	0.51
1:B:324:ARG:HH12	1:B:367:LEU:HD21	1.76	0.50
1:B:359:ILE:CD1	2:E:1:BBV:H5'	2.41	0.50
1:A:274:ASP:OD1	1:A:275:LYS:N	2.45	0.50
1:A:319:ASN:ND2	1:A:359:ILE:HD12	2.27	0.50
1:B:396:THR:HB	1:B:397:PRO:HD2	1.95	0.49
1:C:314:ASP:N	1:C:314:ASP:OD1	2.46	0.49
1:C:291:ALA:C	1:C:292:LEU:HD12	2.33	0.48
1:B:254:MET:SD	1:B:337:THR:HG22	2.54	0.48
1:C:145:PHE:HB2	5:C:406:SO4:O4	2.14	0.48
1:C:266:PRO:HG3	1:C:321:LEU:HD22	1.96	0.48
1:C:244:PHE:O	1:C:291:ALA:HA	2.14	0.48
1:C:303:PHE:O	1:C:370:GLY:HA2	2.13	0.47
1:A:201:VAL:O	1:A:205:GLN:HG3	2.16	0.46
1:A:224:ARG:HD3	1:A:313:GLU:OE1	2.15	0.46
1:B:273:MET:CE	1:B:275:LYS:HE3	2.45	0.46
1:A:180:ILE:HG13	1:A:245:VAL:HB	1.96	0.46
1:C:197:TYR:O	1:C:200:PRO:HD2	2.16	0.45
1:C:245:VAL:HG22	1:C:291:ALA:HB2	1.98	0.45
1:B:167:TYR:CD2	1:B:167:TYR:C	2.90	0.45
1:B:355:ARG:O	1:B:359:ILE:HG23	2.17	0.45
1:C:167:TYR:HB3	1:C:209:TYR:CE2	2.52	0.44
1:C:213:VAL:HG21	1:C:387:TYR:CD1	2.53	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:187:ARG:NH2	3:C:403:UDP:O1A	2.48	0.44
1:A:167:TYR:HB3	1:A:209:TYR:CE2	2.53	0.44
1:A:285:TYR:O	1:A:330:ARG:NH2	2.49	0.44
1:A:174:SER:HB2	1:A:206:GLN:O	2.18	0.44
1:A:268:HIS:HB3	1:A:330:ARG:HG2	1.98	0.44
1:B:313:GLU:OE1	6:B:410:GOL:O3	2.25	0.43
1:C:224:ARG:HD3	1:C:313:GLU:OE1	2.18	0.43
1:B:307:TYR:OH	1:B:316:ASP:HB2	2.18	0.43
1:C:342:ARG:HD2	1:C:342:ARG:HA	1.90	0.43
1:B:153:LEU:O	1:B:157:GLN:HG3	2.18	0.43
1:C:352:ASN:HA	1:C:353:PRO:HD3	1.88	0.43
1:B:153:LEU:HD13	1:B:153:LEU:HA	1.87	0.42
1:A:310:TRP:CD1	6:A:409:GOL:C1	3.01	0.42
1:B:177:LYS:NZ	1:B:240[B]:ASP:OD2	2.51	0.42
1:C:271:VAL:CG2	1:C:336:GLY:HA3	2.50	0.42
1:C:303:PHE:HB3	1:C:304:PRO:HD2	2.02	0.42
1:C:213:VAL:CG2	1:C:387:TYR:CD1	3.03	0.42
1:A:314:ASP:OD1	1:A:314:ASP:N	2.53	0.42
1:C:254:MET:HE3	1:C:339:ARG:HG3	2.02	0.41
1:C:382:GLN:NE2	7:C:522:HOH:O	2.53	0.41
1:A:179:ALA:HB2	1:A:241:TYR:CG	2.56	0.41
1:B:245:VAL:HG22	1:B:291:ALA:HB2	2.02	0.41
1:C:359:ILE:HD12	1:C:359:ILE:HG21	1.85	0.41
1:C:194:TRP:CZ2	1:C:198:LEU:HG	2.54	0.41
1:B:226:LYS:HD3	1:B:394:ILE:HB	2.01	0.41
1:B:273:MET:HE1	1:B:275:LYS:HE3	2.01	0.41
1:A:178:VAL:O	1:A:209:TYR:HA	2.21	0.41
1:B:249:VAL:HG12	3:B:403:UDP:O2'	2.20	0.41
1:C:198:LEU:HD13	1:C:198:LEU:HA	1.71	0.41
1:C:258:ASN:OD1	1:C:331:PRO:HG3	2.21	0.40
1:B:306:ASN:OD1	1:B:358:ARG:NH1	2.55	0.40
1:B:320:ARG:HH11	1:B:320:ARG:HD3	1.74	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

of similar resolution.

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	274/287 (96%)	269 (98%)	5 (2%)	0	100	100
1	B	274/287 (96%)	270 (98%)	4 (2%)	0	100	100
1	C	271/287 (94%)	259 (96%)	12 (4%)	0	100	100
All	All	819/861 (95%)	798 (97%)	21 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	247/252 (98%)	240 (97%)	7 (3%)	43	60
1	B	247/252 (98%)	238 (96%)	9 (4%)	35	49
1	C	244/252 (97%)	235 (96%)	9 (4%)	34	48
All	All	738/756 (98%)	713 (97%)	25 (3%)	38	51

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	147	MET
1	A	153	LEU
1	A	170	ARG
1	A	306	ASN
1	A	347	LYS
1	A	359	ILE
1	A	364	GLU
1	B	127	LEU
1	B	147[A]	MET
1	B	147[B]	MET
1	B	153	LEU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	170	ARG
1	B	306	ASN
1	B	345	ARG
1	B	359	ILE
1	B	394	ILE
1	C	147	MET
1	C	170	ARG
1	C	219	ASP
1	C	233	GLN
1	C	342	ARG
1	C	347	LYS
1	C	359	ILE
1	C	373	SER
1	C	398	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	257	HIS
1	A	306	ASN
1	A	319	ASN
1	B	206	GLN
1	B	257	HIS
1	B	306	ASN
1	B	319	ASN
1	C	206	GLN

### 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](#)

6 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	BBV	D	1	2	23,23,23	0.88	1 (4%)	30,31,31	1.29	4 (13%)
2	NAG	D	2	2	14,14,15	0.46	0	17,19,21	1.05	1 (5%)
2	BBV	E	1	2	23,23,23	1.06	1 (4%)	30,31,31	1.18	4 (13%)
2	NAG	E	2	2	14,14,15	0.90	0	17,19,21	1.28	2 (11%)
2	BBV	F	1	2	23,23,23	0.88	1 (4%)	30,31,31	1.89	4 (13%)
2	NAG	F	2	2	14,14,15	0.68	0	17,19,21	0.91	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	BBV	D	1	2	-	1/11/31/31	0/2/2/2
2	NAG	D	2	2	-	1/6/23/26	0/1/1/1
2	BBV	E	1	2	-	1/11/31/31	0/2/2/2
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1
2	BBV	F	1	2	-	2/11/31/31	0/2/2/2
2	NAG	F	2	2	-	0/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	1	BBV	O1-C7'	-3.68	1.32	1.42
2	F	1	BBV	O1-C7'	-3.48	1.33	1.42
2	D	1	BBV	O1-C7'	-3.19	1.34	1.42

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	1	BBV	C7'-O1-C1	6.52	124.89	113.31
2	F	1	BBV	C3-C2-N2	4.39	118.90	110.62
2	D	1	BBV	C7'-O1-C1	4.03	120.47	113.31

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	1	BBV	C1-C2-N2	-3.92	104.26	111.00
2	E	2	NAG	C2-N2-C7	-3.17	118.39	122.90
2	F	1	BBV	C1-O5-C5	3.15	119.87	113.69
2	E	1	BBV	C3-C2-N2	2.80	115.90	110.62
2	D	2	NAG	C1-O5-C5	2.64	115.78	112.19
2	D	1	BBV	O3-C3-C2	-2.39	104.83	109.66
2	F	2	NAG	C1-O5-C5	2.38	115.41	112.19
2	E	1	BBV	O5-C5-C4	-2.38	105.38	109.69
2	E	1	BBV	O3-C3-C4	-2.22	105.22	110.35
2	E	1	BBV	C4-C3-C2	-2.15	107.20	110.34
2	E	2	NAG	C3-C4-C5	-2.09	106.51	110.24
2	D	1	BBV	O5-C5-C6	2.05	111.52	106.44
2	D	1	BBV	C3'-C1'-C2'	2.01	121.33	118.17

There are no chirality outliers.

All (5) torsion outliers are listed below:

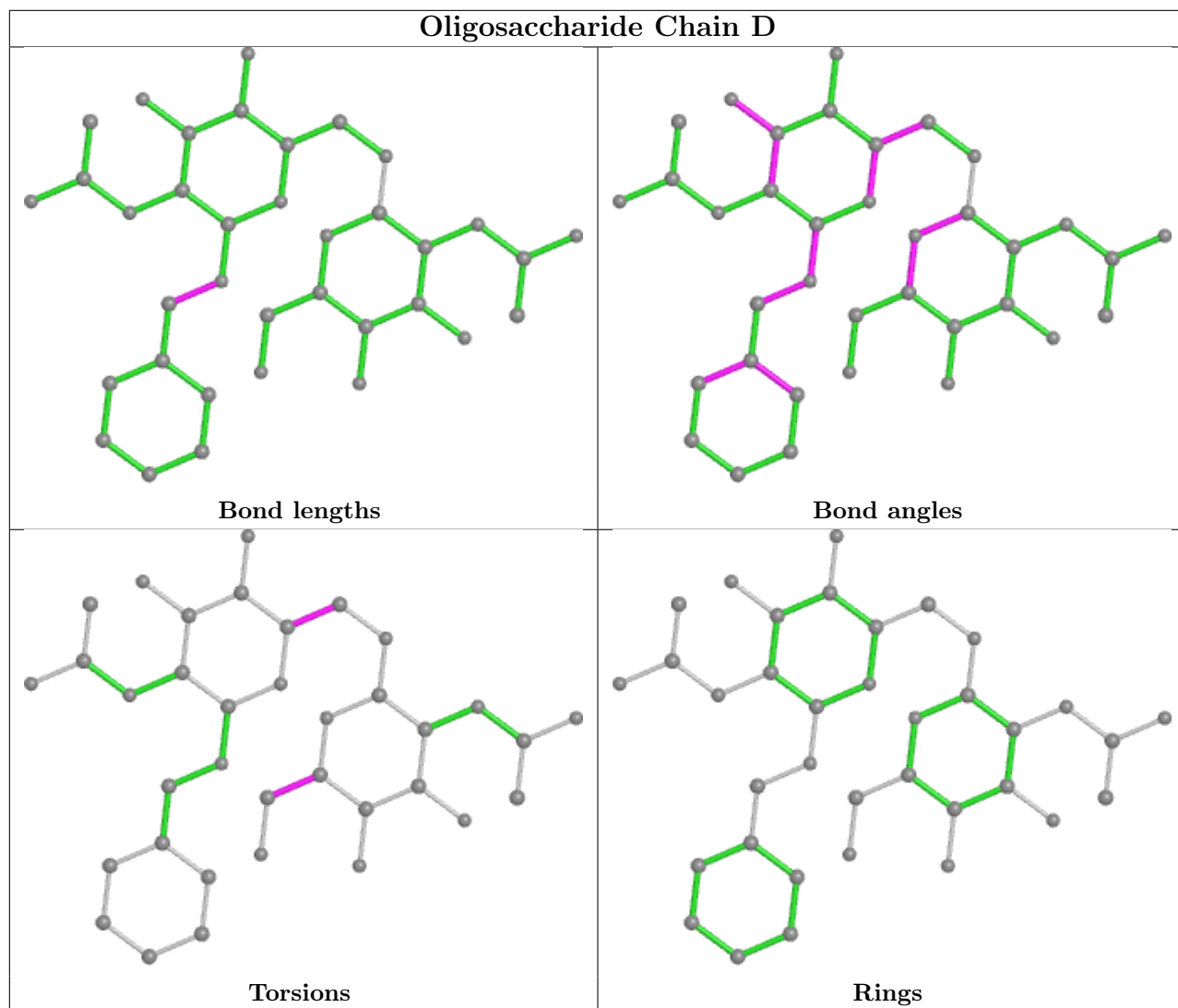
Mol	Chain	Res	Type	Atoms
2	F	1	BBV	O5-C1-O1-C7'
2	D	1	BBV	C4-C5-C6-O6
2	E	1	BBV	C2-C1-O1-C7'
2	D	2	NAG	C4-C5-C6-O6
2	F	1	BBV	C4-C5-C6-O6

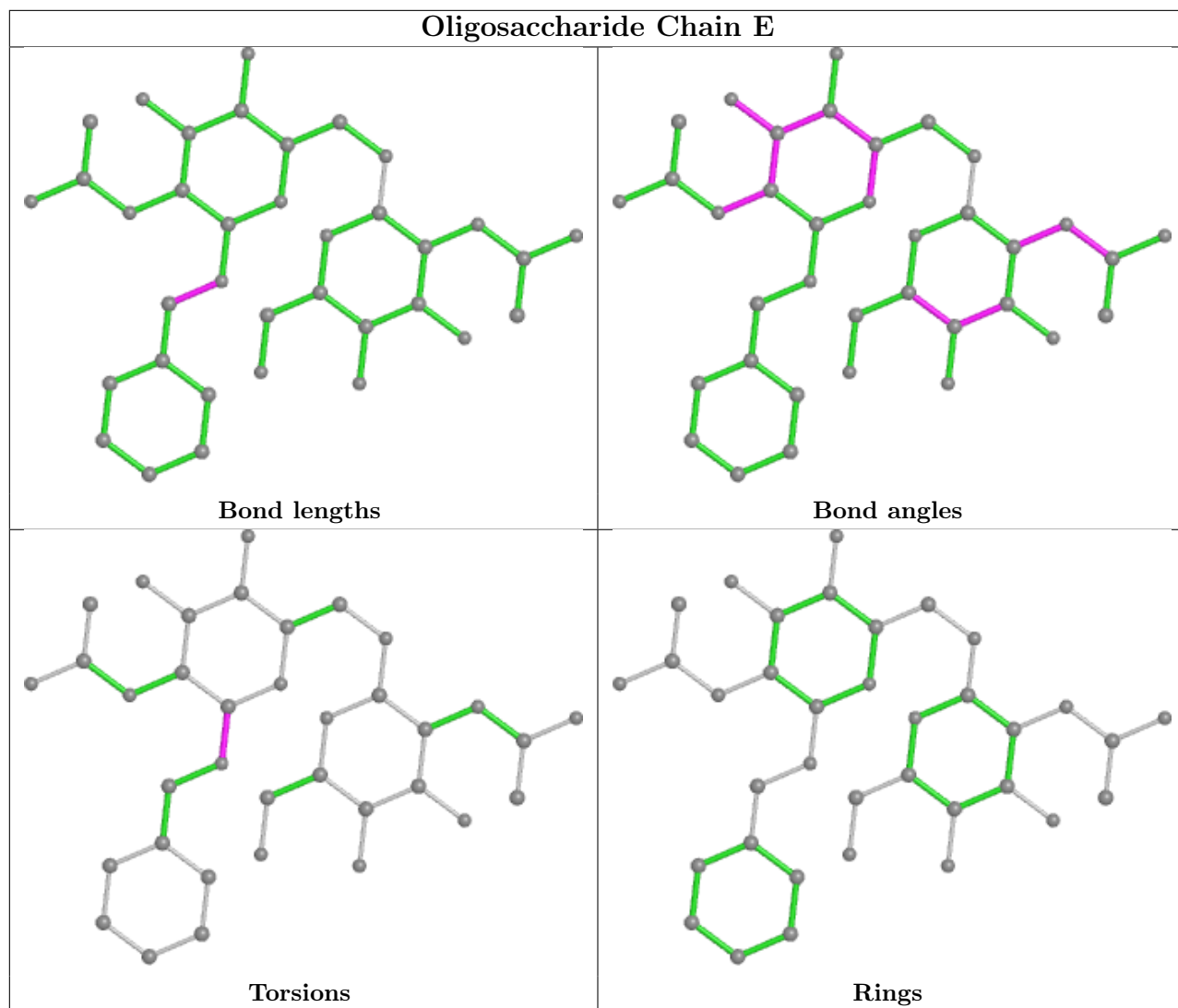
There are no ring outliers.

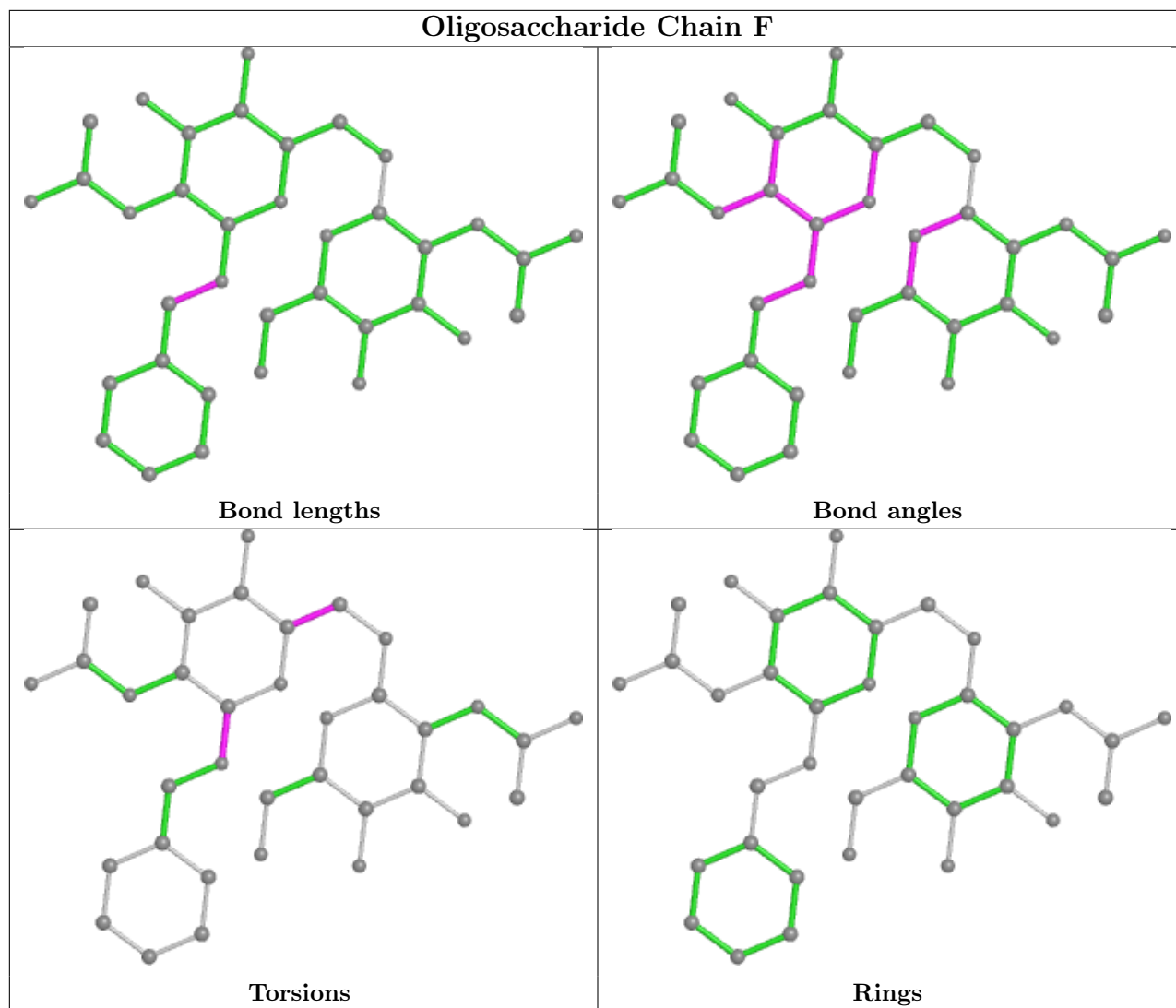
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	1	BBV	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 20 ligands modelled in this entry, 3 are monoatomic - leaving 17 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	SO4	A	408	-	4,4,4	0.53	0	6,6,6	0.81	0
5	SO4	B	408	-	4,4,4	0.23	0	6,6,6	0.25	0
6	GOL	B	410	-	5,5,5	0.66	0	5,5,5	1.42	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	SO4	A	405	-	4,4,4	0.26	0	6,6,6	0.63	0
3	UDP	B	403	4	24,26,26	0.99	1 (4%)	37,40,40	2.11	5 (13%)
3	UDP	C	403	4	24,26,26	0.91	1 (4%)	37,40,40	1.91	8 (21%)
5	SO4	A	407	-	4,4,4	0.20	0	6,6,6	0.43	0
5	SO4	B	409	-	4,4,4	0.16	0	6,6,6	0.55	0
6	GOL	A	409	-	5,5,5	0.37	0	5,5,5	0.78	0
5	SO4	B	405	-	4,4,4	0.40	0	6,6,6	0.57	0
3	UDP	A	403	4	24,26,26	1.09	1 (4%)	37,40,40	1.61	6 (16%)
5	SO4	B	406	-	4,4,4	0.44	0	6,6,6	0.64	0
5	SO4	B	407	-	4,4,4	0.17	0	6,6,6	0.26	0
5	SO4	A	406	-	4,4,4	0.15	0	6,6,6	0.47	0
5	SO4	C	405	-	4,4,4	0.17	0	6,6,6	0.31	0
6	GOL	C	407	-	5,5,5	0.40	0	5,5,5	0.45	0
5	SO4	C	406	-	4,4,4	0.17	0	6,6,6	0.26	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	GOL	B	410	-	-	0/4/4/4	-
3	UDP	B	403	4	-	1/16/32/32	0/2/2/2
6	GOL	A	409	-	-	4/4/4/4	-
3	UDP	C	403	4	-	2/16/32/32	0/2/2/2
3	UDP	A	403	4	-	1/16/32/32	0/2/2/2
6	GOL	C	407	-	-	1/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	403	UDP	C6-C5	3.12	1.42	1.35
3	A	403	UDP	C6-C5	2.63	1.41	1.35
3	C	403	UDP	C5-C4	-2.28	1.38	1.43

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	403	UDP	C4-N3-C2	-6.79	117.62	126.58
3	B	403	UDP	N3-C2-N1	6.62	123.68	114.89
3	C	403	UDP	C4-N3-C2	-5.57	119.24	126.58

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	403	UDP	N3-C2-N1	4.65	121.06	114.89
3	A	403	UDP	C4-N3-C2	-4.20	121.03	126.58
3	A	403	UDP	PA-O3A-PB	-4.18	118.49	132.83
3	B	403	UDP	C5-C4-N3	3.87	120.63	114.84
3	C	403	UDP	C5-C4-N3	3.80	120.53	114.84
3	A	403	UDP	N3-C2-N1	3.60	119.67	114.89
3	B	403	UDP	PA-O3A-PB	-3.56	120.61	132.83
3	C	403	UDP	PA-O3A-PB	-3.49	120.84	132.83
3	C	403	UDP	O4-C4-C5	-3.37	119.24	125.16
3	B	403	UDP	O2-C2-N3	-3.14	115.64	121.50
3	A	403	UDP	C5-C4-N3	2.76	118.97	114.84
3	C	403	UDP	O4'-C4'-C5'	2.55	117.76	109.37
3	C	403	UDP	O4'-C1'-N1	2.31	113.66	108.36
3	C	403	UDP	C1'-N1-C2	2.23	121.61	117.57
3	A	403	UDP	C6-N1-C2	-2.16	118.23	120.99
3	A	403	UDP	O4'-C1'-N1	2.03	113.01	108.36

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	409	GOL	O1-C1-C2-C3
6	A	409	GOL	C1-C2-C3-O3
6	A	409	GOL	O1-C1-C2-O2
6	A	409	GOL	O2-C2-C3-O3
3	B	403	UDP	C4'-C5'-O5'-PA
3	A	403	UDP	C4'-C5'-O5'-PA
6	C	407	GOL	C1-C2-C3-O3
3	C	403	UDP	C4'-C5'-O5'-PA
3	C	403	UDP	C5'-O5'-PA-O1A

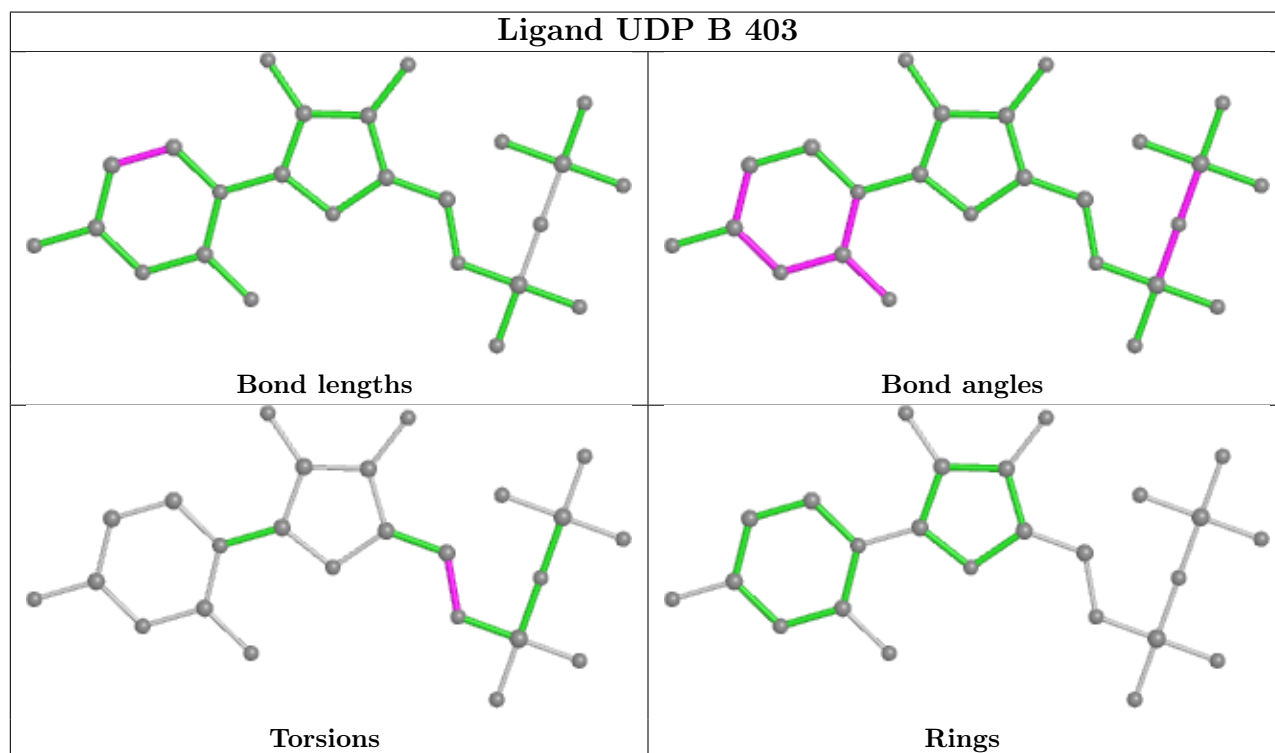
There are no ring outliers.

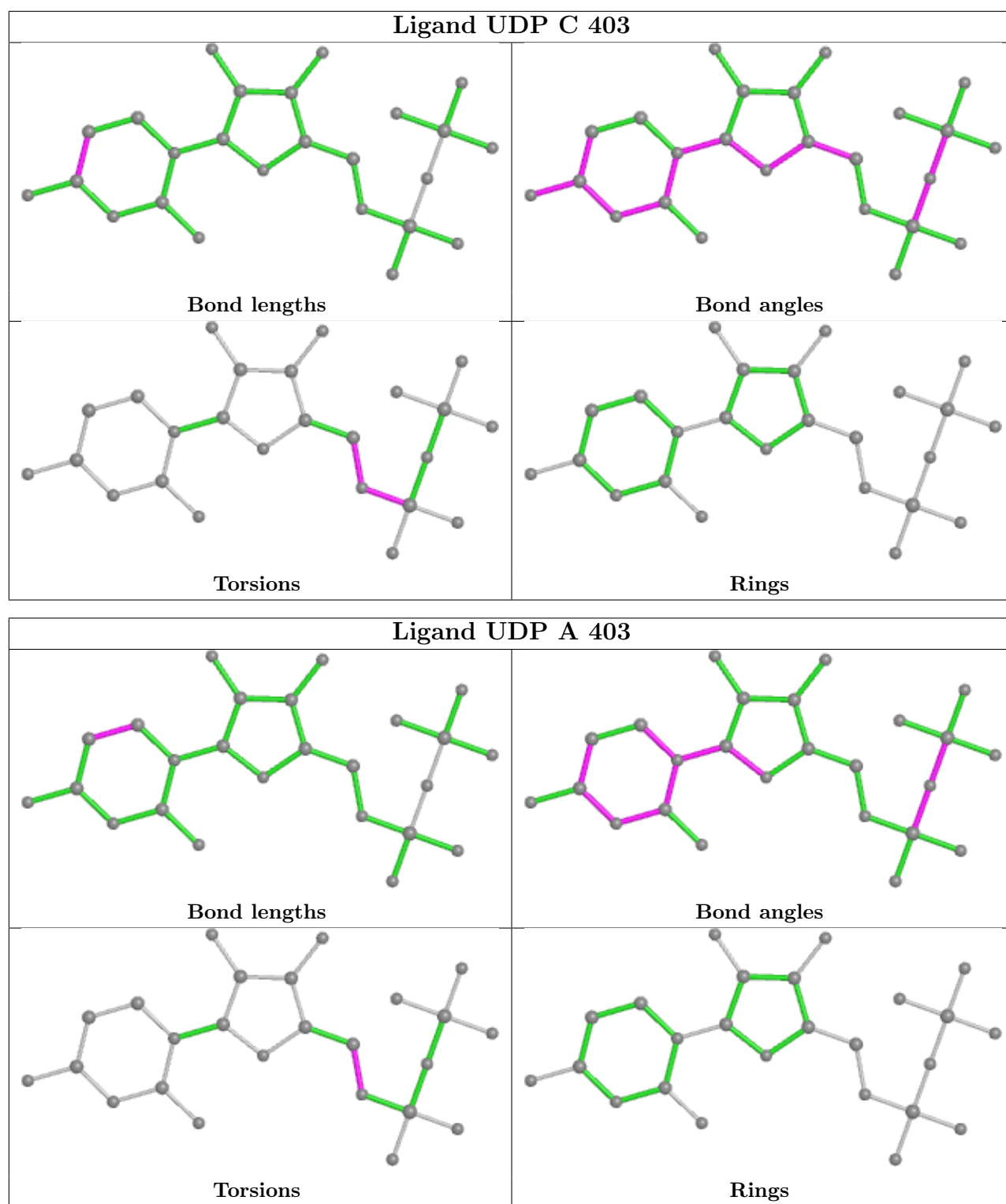
5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	410	GOL	2	0
3	B	403	UDP	1	0
3	C	403	UDP	2	0
6	A	409	GOL	2	0
5	C	406	SO4	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 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

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	273/287 (95%)	-0.31	4 (1%) 73 79	28, 37, 57, 72	0
1	B	273/287 (95%)	-0.30	1 (0%) 92 95	26, 34, 48, 62	0
1	C	273/287 (95%)	-0.04	7 (2%) 56 63	33, 48, 71, 82	0
All	All	819/861 (95%)	-0.22	12 (1%) 73 79	26, 39, 63, 82	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	126	SER	4.7
1	C	126	SER	4.3
1	C	398	SER	3.4
1	C	347	LYS	3.1
1	A	353	PRO	2.4
1	C	357	ASP	2.3
1	C	306	ASN	2.3
1	C	219	ASP	2.3
1	A	347	LYS	2.2
1	A	345	ARG	2.2
1	C	148	PRO	2.2
1	A	126	SER	2.0

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

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

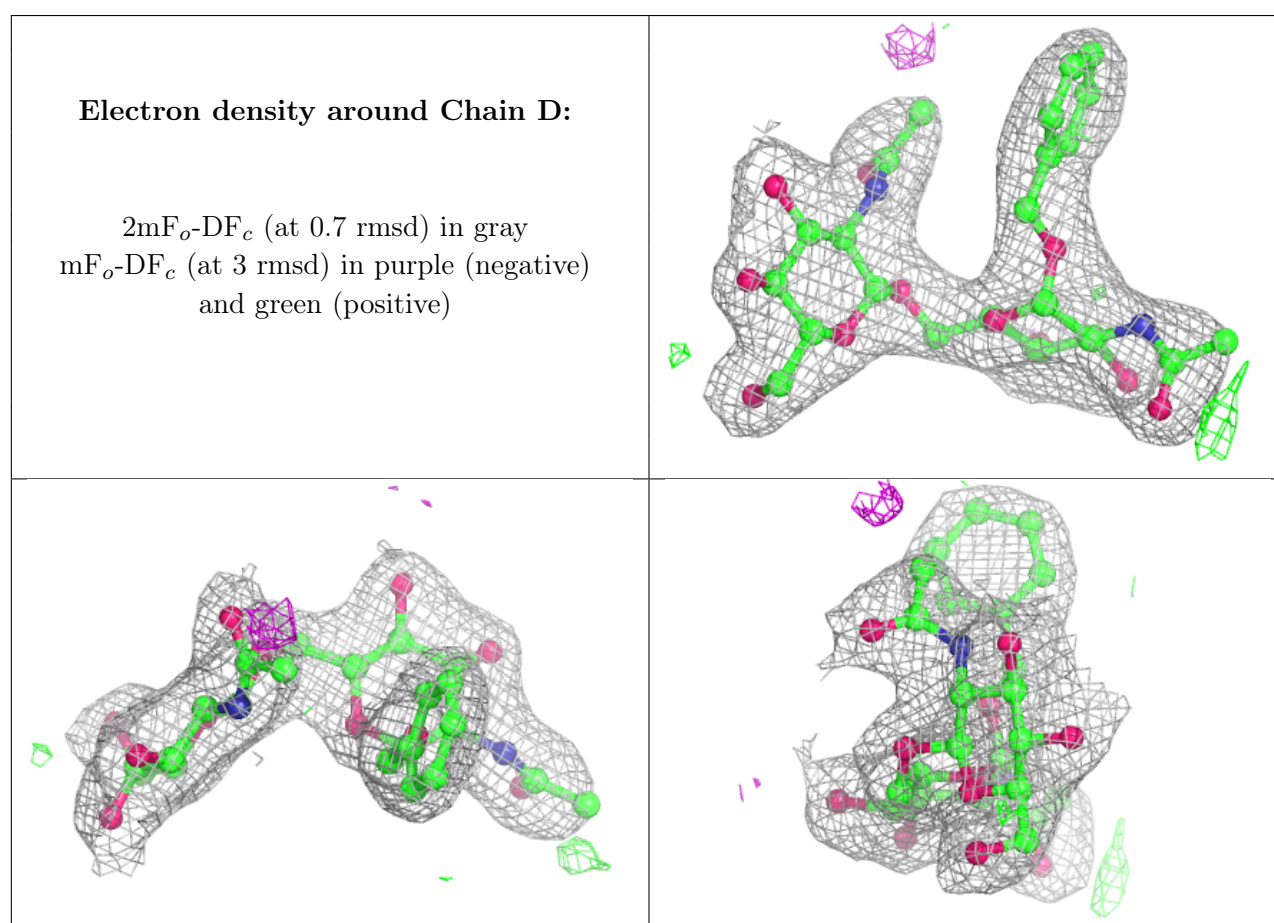
### 6.3 Carbohydrates [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

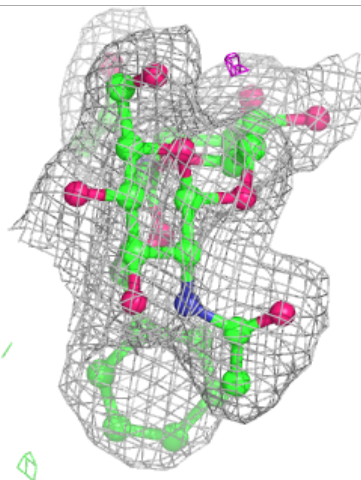
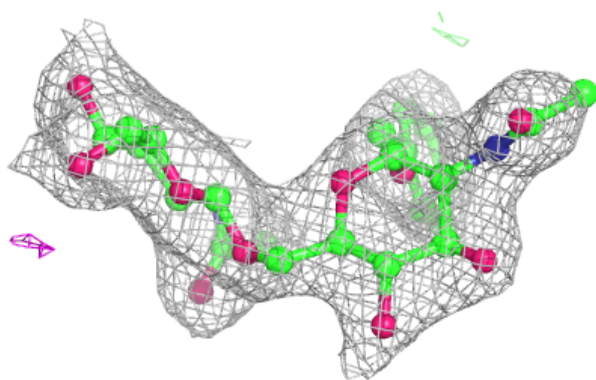
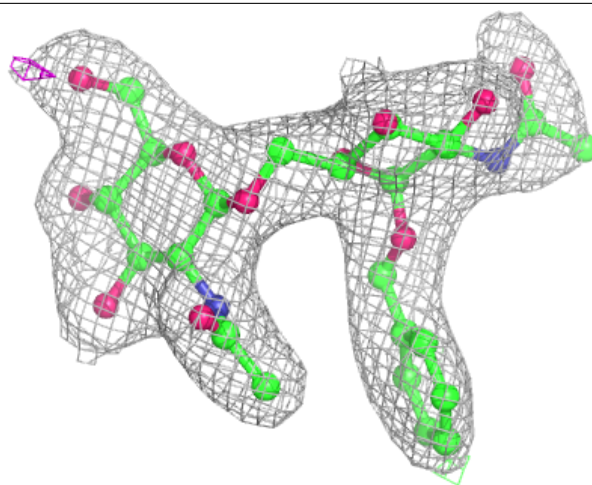
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	BBV	F	1	22/22	0.93	0.15	67,78,85,86	0
2	NAG	F	2	14/15	0.94	0.16	57,60,63,64	0
2	BBV	E	1	22/22	0.96	0.15	41,46,56,58	0
2	NAG	E	2	14/15	0.96	0.11	29,34,36,37	0
2	BBV	D	1	22/22	0.97	0.10	38,44,54,57	0
2	NAG	D	2	14/15	0.97	0.09	34,38,42,43	0

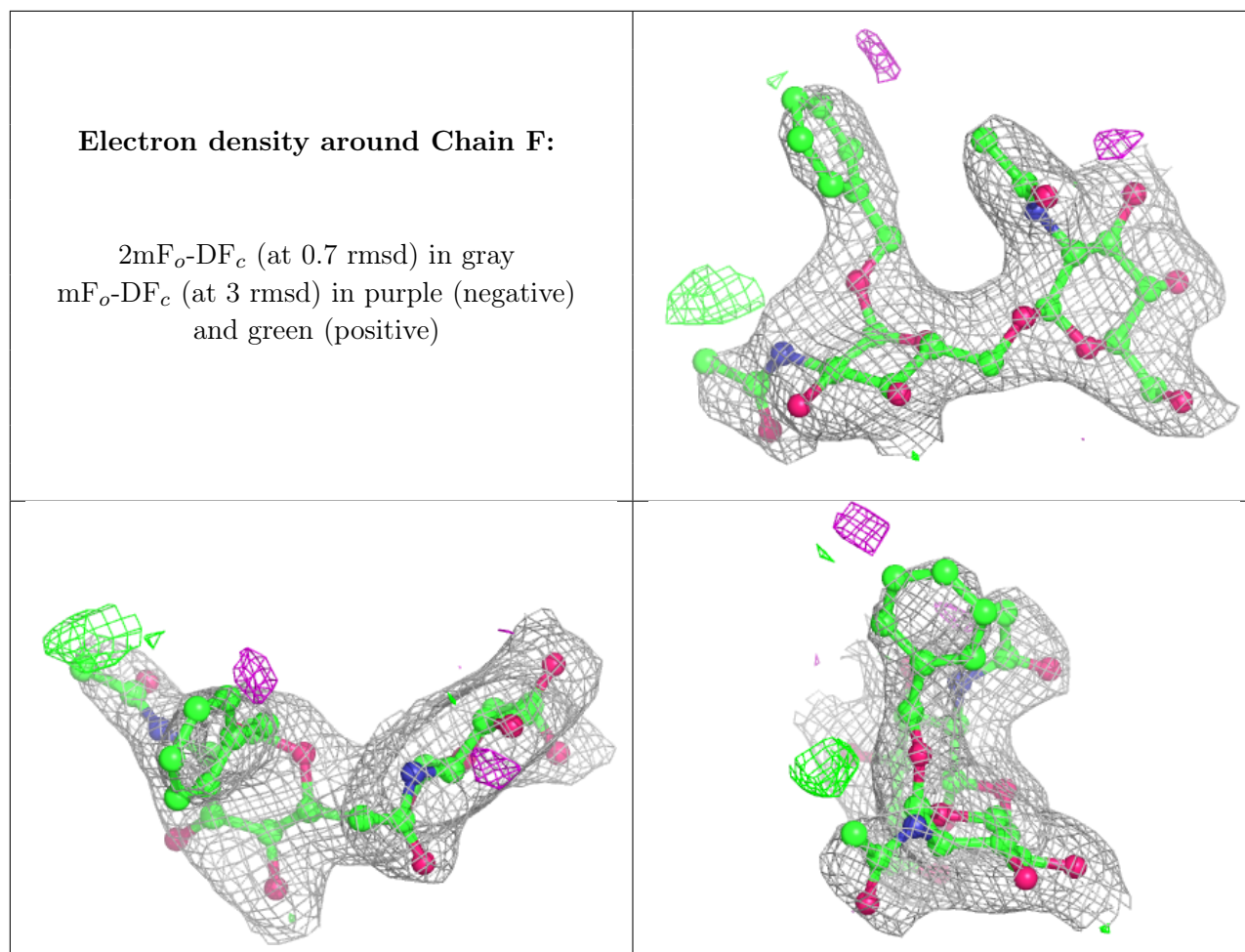
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**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)





## 6.4 Ligands [i](#)

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

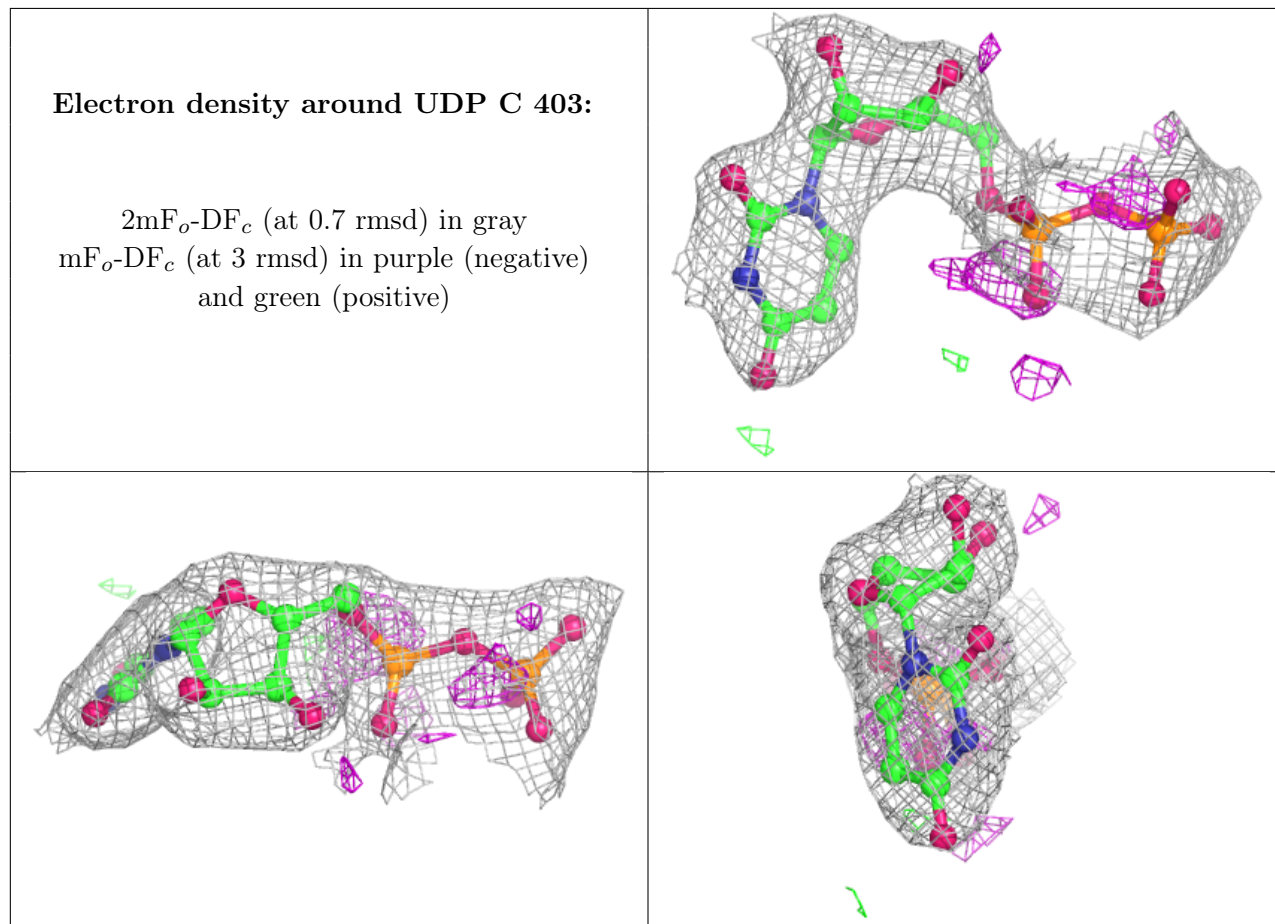
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	SO4	A	408	5/5	0.70	0.30	66,70,77,81	0
5	SO4	A	407	5/5	0.83	0.27	86,92,93,94	0
3	UDP	C	403	25/25	0.88	0.16	52,57,90,93	0
3	UDP	B	403	25/25	0.89	0.15	26,31,79,82	0
5	SO4	B	406	5/5	0.89	0.16	66,70,73,75	0
5	SO4	B	409	5/5	0.90	0.28	85,86,88,90	0
3	UDP	A	403	25/25	0.91	0.14	33,40,74,77	0
5	SO4	B	405	5/5	0.91	0.36	60,65,71,72	0
6	GOL	C	407	6/6	0.91	0.15	62,65,67,68	0
6	GOL	B	410	6/6	0.92	0.15	39,51,52,52	0
5	SO4	A	406	5/5	0.92	0.16	85,87,91,92	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	SO4	A	405	5/5	0.93	0.10	67,67,72,73	0
5	SO4	C	406	5/5	0.93	0.12	94,96,97,97	0
5	SO4	B	407	5/5	0.95	0.17	87,88,88,89	0
6	GOL	A	409	6/6	0.96	0.16	42,51,56,60	0
5	SO4	C	405	5/5	0.96	0.19	75,77,77,81	0
5	SO4	B	408	5/5	0.96	0.11	62,64,65,70	0
4	MN	A	404	1/1	0.98	0.07	48,48,48,48	0
4	MN	C	404	1/1	0.98	0.07	60,60,60,60	0
4	MN	B	404	1/1	0.99	0.06	41,41,41,41	0

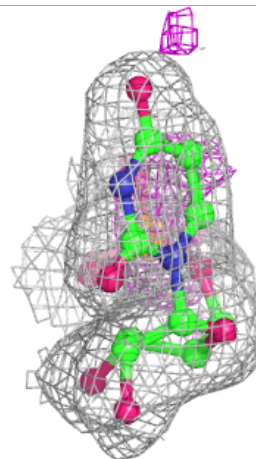
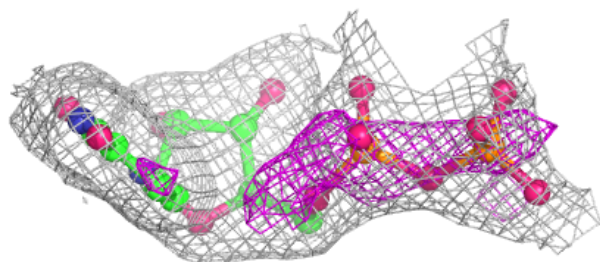
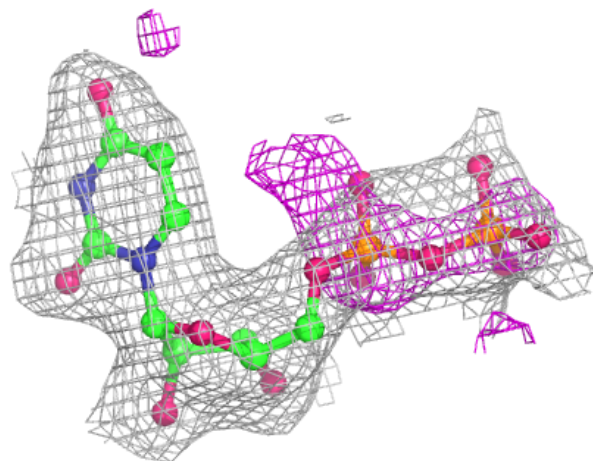
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

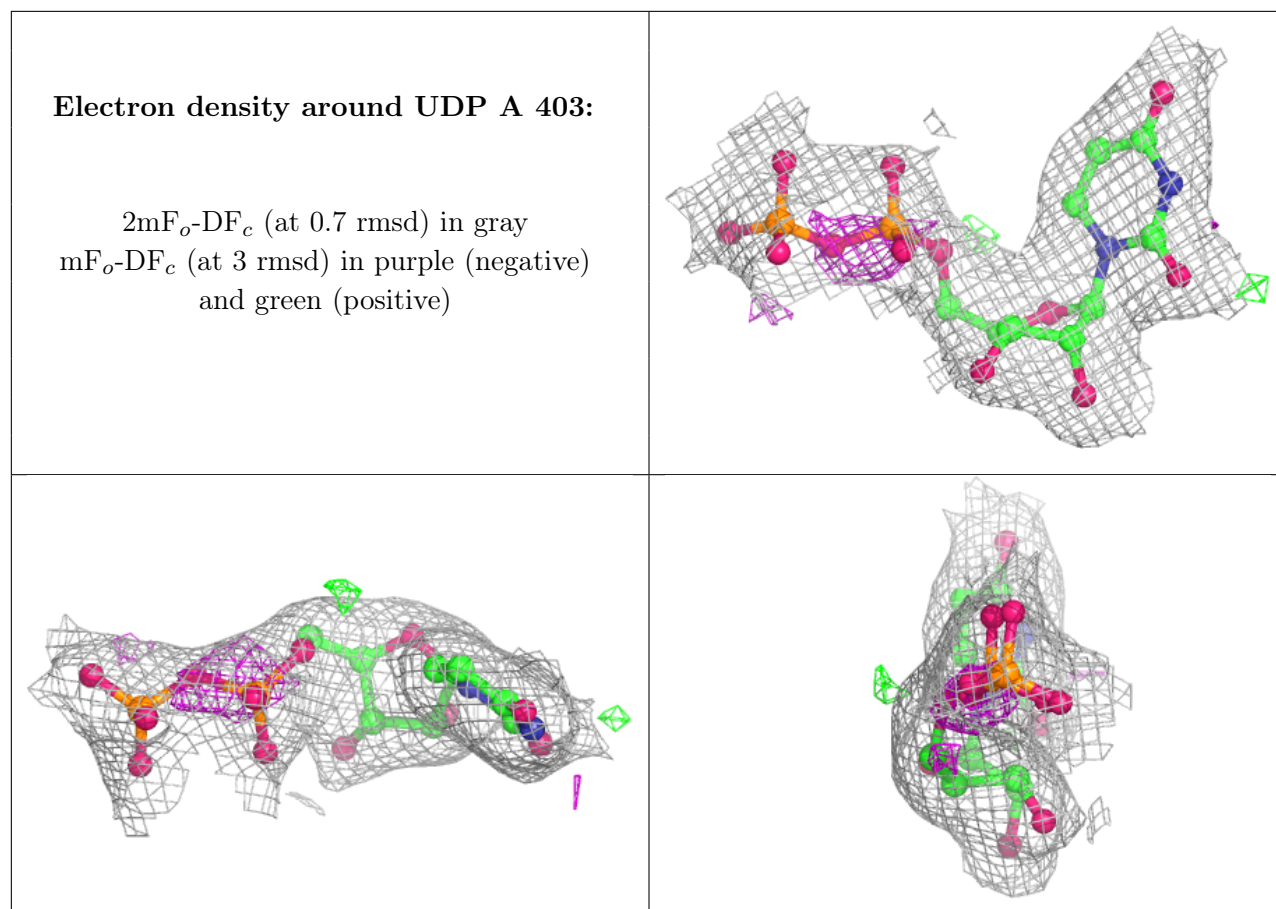




**Electron density around UDP B 403:**

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





## 6.5 Other polymers [i](#)

There are no such residues in this entry.