



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 31, 2023 – 10:45 PM EDT

PDB ID : 2CHB  
Title : CHOLERA TOXIN B-PENTAMER COMPLEXED WITH GM1 PENTASACCHARIDE  
Authors : Merritt, E.A.; Hol, W.G.J.  
Deposited on : 1997-06-03  
Resolution : 2.00 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 : 2.34  
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) : 2.34

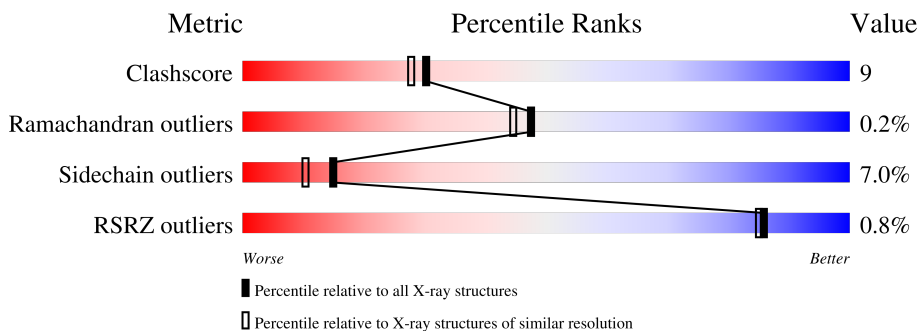
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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.



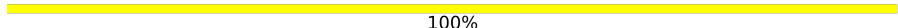
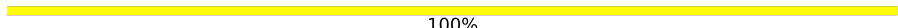
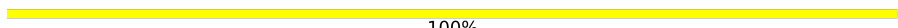
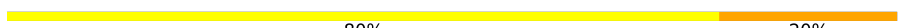
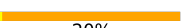
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	D	104	 75% 20% . .
1	E	104	 2% 72% 21% 6% .
1	F	104	 2% 77% 21% . .
1	G	104	 2% 77% 20% . . .
1	H	104	 75% 20% . .
2	A	4	 75% 25%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	C	4	 100%
3	B	5	 100%
3	I	5	 100%
3	J	5	 80%  20%

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 4774 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.

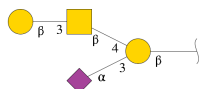
- Molecule 1 is a protein called CHOLERA TOXIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	D	103	Total 814	C 511	N 142	O 156	S 5	0	0	0
1	E	103	Total 814	C 511	N 142	O 156	S 5	0	0	0
1	F	103	Total 814	C 511	N 142	O 156	S 5	0	0	0
1	G	103	Total 814	C 511	N 142	O 156	S 5	0	0	0
1	H	103	Total 814	C 511	N 142	O 156	S 5	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

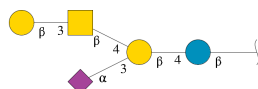
Chain	Residue	Modelled	Actual	Comment	Reference
D	18	HIS	TYR	conflict	UNP P01556
D	47	THR	ILE	conflict	UNP P01556
E	18	HIS	TYR	conflict	UNP P01556
E	47	THR	ILE	conflict	UNP P01556
F	18	HIS	TYR	conflict	UNP P01556
F	47	THR	ILE	conflict	UNP P01556
G	18	HIS	TYR	conflict	UNP P01556
G	47	THR	ILE	conflict	UNP P01556
H	18	HIS	TYR	conflict	UNP P01556
H	47	THR	ILE	conflict	UNP P01556

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	A	4	Total	C	N	O	0	0	0
			56	31	2	23			
2	C	4	Total	C	N	O	0	0	0
			56	31	2	23			

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	B	5	Total	C	N	O	0	0	0
			68	37	2	29			
3	I	5	Total	C	N	O	0	0	0
			68	37	2	29			
3	J	5	Total	C	N	O	0	0	0
			68	37	2	29			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	78	Total	O	0	0
			78	78		
4	E	77	Total	O	0	0
			77	77		
4	F	47	Total	O	0	0
			47	47		
4	G	90	Total	O	0	0
			90	90		
4	H	96	Total	O	0	0
			96	96		

### 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: CHOLERA TOXIN

Chain D: 




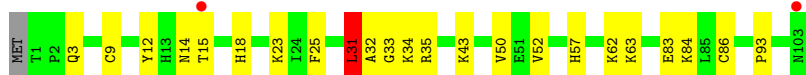
- Molecule 1: CHOLERA TOXIN

Chain E: 




- Molecule 1: CHOLERA TOXIN

Chain F: 



- Molecule 1: CHOLERA TOXIN

Chain G: 




- Molecule 1: CHOLERA TOXIN

Chain H: 

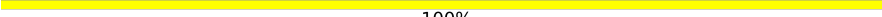


- Molecule 2: beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-galactopyranose-(1-4)-[N-acetyl-alpha-neuraminic acid-(2-3)]beta-D-galactopyranose

Chain A:  75% 25%

GAL1  
NGA2  
GAL3  
SIA4

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

Chain C:  100%

GAL1  
NGA2  
GAL3  
SIA4

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

Chain B:  100%


BGC1  
GAL2  
NGA3  
GAL4  
SIA5

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

Chain I:  100%

BGC1  
GAL2  
NGA3  
GAL4  
SIA5

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

Chain J:  80% 20%

BGC1  
GAL2  
NGA3  
GAL4  
SIA5

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.90Å 67.58Å 80.47Å 90.00° 105.69° 90.00°	Depositor
Resolution (Å)	10.00 – 2.00 35.41 – 1.78	Depositor EDS
% Data completeness (in resolution range)	81.4 (10.00-2.00) 84.3 (35.41-1.78)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.50 (at 1.78Å)	Xtrriage
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	0.171 , (Not available) 0.184 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.5	Xtrriage
Anisotropy	0.506	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 85.9	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	4774	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.37% 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: GAL, BGC, SIA, NGA

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	D	0.53	0/828	0.77	2/1118 (0.2%)
1	E	0.51	0/828	0.76	2/1118 (0.2%)
1	F	0.53	0/828	0.79	2/1118 (0.2%)
1	G	0.61	1/828 (0.1%)	0.84	3/1118 (0.3%)
1	H	0.63	1/828 (0.1%)	0.80	2/1118 (0.2%)
All	All	0.56	2/4140 (0.0%)	0.79	11/5590 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	68	MET	CG-SD	9.19	2.05	1.81
1	G	68	MET	CG-SD	-9.18	1.57	1.81

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	31	LEU	CA-CB-CG	7.38	132.28	115.30
1	F	31	LEU	CA-CB-CG	7.10	131.63	115.30
1	G	31	LEU	CA-CB-CG	6.57	130.41	115.30
1	H	31	LEU	CA-CB-CG	6.38	129.97	115.30
1	E	50	VAL	N-CA-C	-6.24	94.16	111.00
1	H	50	VAL	N-CA-C	-6.08	94.58	111.00
1	D	50	VAL	N-CA-C	-5.96	94.91	111.00
1	G	99	ILE	CB-CA-C	-5.60	100.39	111.60
1	G	50	VAL	N-CA-C	-5.58	95.93	111.00
1	E	73	ARG	NE-CZ-NH2	-5.44	117.58	120.30
1	F	50	VAL	N-CA-C	-5.24	96.85	111.00

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	D	814	0	817	21	0
1	E	814	0	817	20	0
1	F	814	0	817	22	0
1	G	814	0	817	16	0
1	H	814	0	817	18	0
2	A	56	0	47	2	0
2	C	56	0	47	0	0
3	B	68	0	57	0	0
3	I	68	0	58	0	0
3	J	68	0	58	1	0
4	D	78	0	0	2	0
4	E	77	0	0	4	0
4	F	47	0	0	2	0
4	G	90	0	0	1	0
4	H	96	0	0	4	0
All	All	4774	0	4352	81	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 9.

All (81) 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:H:68:MET:SD	1:H:68:MET:CG	2.05	1.45
1:F:9:CYS:SG	1:F:15:THR:HG22	2.14	0.88
1:E:3:GLN:HG2	1:F:93:PRO:HG3	1.73	0.69
1:D:77:LEU:HD11	1:H:74:ILE:HD13	1.75	0.69
1:D:73:ARG:HD2	1:H:74:ILE:HD11	1.76	0.67
1:F:15:THR:CG2	1:F:86:CYS:SG	2.83	0.67
1:G:85:LEU:HD23	1:G:99:ILE:HG23	1.77	0.66
1:D:2:PRO:HG2	1:E:35:ARG:HD3	1.78	0.65
1:F:25:PHE:HB2	1:F:43:LYS:HD2	1.80	0.64
1:F:83:GLU:HG2	1:F:84:LYS:HG3	1.80	0.63
1:G:63:LYS:HD3	1:G:64:ALA:N	2.14	0.63
1:E:83:GLU:HG2	1:E:84:LYS:HG3	1.82	0.62

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:25:PHE:HD1	1:F:43:LYS:HZ3	1.49	0.61
1:F:15:THR:HG21	1:F:86:CYS:SG	2.41	0.61
1:F:9:CYS:SG	1:F:15:THR:CG2	2.88	0.60
1:G:99:ILE:CD1	1:H:29:GLU:HB3	2.32	0.60
1:H:1:THR:HB	1:H:2:PRO:CD	2.32	0.59
1:H:56:GLN:HG3	3:J:4:GAL:O4	2.02	0.59
1:H:81:LYS:HD3	4:H:162:HOH:O	2.01	0.59
1:G:85:LEU:CD2	1:G:99:ILE:HG23	2.33	0.59
1:D:34:LYS:HD3	4:H:182:HOH:O	2.02	0.58
1:E:16:GLN:NE2	1:E:89:ASN:HD22	2.02	0.58
1:H:89:ASN:HB3	4:H:184:HOH:O	2.02	0.58
1:G:4:ASN:HD22	1:G:6:THR:H	1.52	0.57
1:E:63:LYS:HG3	4:E:118:HOH:O	2.03	0.57
1:D:93:PRO:HA	4:D:124:HOH:O	2.06	0.56
1:G:4:ASN:HD21	1:G:6:THR:HB	1.70	0.56
1:H:68:MET:CG	1:H:68:MET:CE	2.83	0.55
1:G:15:THR:HA	1:G:87:VAL:O	2.07	0.55
1:G:102:ALA:O	1:G:103:ASN:HB2	2.08	0.54
1:F:63:LYS:HG3	4:G:144:HOH:O	2.08	0.54
1:E:1:THR:HB	1:E:2:PRO:HD2	1.89	0.53
1:E:14:ASN:HD22	1:E:14:ASN:N	2.05	0.53
1:F:43:LYS:CD	1:F:43:LYS:H	2.22	0.53
1:D:49:GLN:HB3	1:D:93:PRO:HG2	1.91	0.53
1:E:94:HIS:HE1	4:E:148:HOH:O	1.93	0.52
1:E:97:ALA:HA	1:F:31:LEU:HD22	1.91	0.51
1:F:14:ASN:HB2	4:F:154:HOH:O	2.11	0.51
1:G:3:GLN:HG3	1:H:47:THR:HG21	1.92	0.51
1:H:26:SER:OG	1:H:41:THR:HB	2.11	0.51
1:F:43:LYS:HD2	1:F:43:LYS:H	1.76	0.51
1:D:1:THR:HG23	4:E:130:HOH:O	2.11	0.51
1:E:22:ASP:HA	1:E:81:LYS:HE2	1.93	0.50
1:E:44:ASN:HD21	1:E:46:ALA:HB2	1.75	0.49
1:F:23:LYS:HZ2	1:F:43:LYS:HE3	1.77	0.49
1:E:56:GLN:HG2	4:E:165:HOH:O	2.13	0.49
1:G:14:ASN:O	1:G:88:TRP:HA	2.13	0.49
1:D:12:TYR:CZ	1:E:32:ALA:HB1	2.49	0.48
1:D:2:PRO:HG3	1:D:11:GLU:OE2	2.13	0.48
1:D:26:SER:HB2	1:D:41:THR:OG1	2.14	0.48
1:G:4:ASN:ND2	1:G:6:THR:HB	2.29	0.48
1:H:68:MET:SD	1:H:68:MET:CB	2.99	0.47
1:D:77:LEU:CD1	1:H:74:ILE:HD13	2.45	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:74:ILE:HD12	1:E:74:ILE:HA	1.80	0.46
1:D:33:GLY:O	1:D:34:LYS:HB2	2.15	0.46
1:G:99:ILE:HD11	1:H:29:GLU:HB3	1.98	0.45
1:G:63:LYS:HE3	1:H:66:GLU:OE1	2.18	0.44
1:F:33:GLY:O	1:F:34:LYS:HB2	2.18	0.44
1:D:28:THR:O	1:D:38:ALA:HA	2.18	0.44
1:E:12:TYR:OH	1:F:35:ARG:HB2	2.17	0.44
1:H:33:GLY:O	1:H:34:LYS:HB2	2.18	0.44
1:G:34:LYS:O	1:G:35:ARG:HD2	2.17	0.43
1:H:25:PHE:HB3	1:H:41:THR:HG22	2.00	0.43
1:D:94:HIS:N	4:D:124:HOH:O	2.48	0.43
1:G:33:GLY:O	1:G:34:LYS:HB2	2.18	0.43
1:D:13:HIS:HB2	2:A:4:SIA:O1B	2.19	0.43
1:D:87:VAL:CG1	1:D:94:HIS:HB3	2.49	0.42
1:E:68:MET:O	1:E:72:LEU:HG	2.19	0.42
1:D:1:THR:HB	1:E:92:THR:O	2.20	0.42
1:E:12:TYR:CZ	1:F:32:ALA:HB1	2.55	0.42
1:D:92:THR:HA	1:D:93:PRO:C	2.41	0.41
1:E:41:THR:HA	1:E:46:ALA:O	2.20	0.41
1:F:18:HIS:O	1:F:84:LYS:HA	2.20	0.41
1:H:100:SER:HB2	4:H:165:HOH:O	2.20	0.41
1:D:12:TYR:CZ	2:A:4:SIA:H112	2.56	0.41
1:F:52:VAL:HG23	4:F:139:HOH:O	2.20	0.41
1:F:12:TYR:CZ	1:G:32:ALA:HB1	2.56	0.41
1:F:34:LYS:O	1:F:35:ARG:HD2	2.21	0.41
1:F:57:HIS:HB2	1:F:62:LYS:HE2	2.02	0.40
1:D:2:PRO:CG	1:E:35:ARG:HD3	2.50	0.40
1:D:83:GLU:HG2	1:D:84:LYS:HG3	2.02	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	D	101/104 (97%)	98 (97%)	2 (2%)	1 (1%)	15	9
1	E	101/104 (97%)	100 (99%)	1 (1%)	0	100	100
1	F	101/104 (97%)	100 (99%)	1 (1%)	0	100	100
1	G	101/104 (97%)	100 (99%)	1 (1%)	0	100	100
1	H	101/104 (97%)	100 (99%)	1 (1%)	0	100	100
All	All	505/520 (97%)	498 (99%)	6 (1%)	1 (0%)	47	44

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	56	GLN

### 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	D	89/90 (99%)	83 (93%)	6 (7%)	16	11
1	E	89/90 (99%)	81 (91%)	8 (9%)	9	6
1	F	89/90 (99%)	87 (98%)	2 (2%)	52	55
1	G	89/90 (99%)	84 (94%)	5 (6%)	21	17
1	H	89/90 (99%)	79 (89%)	10 (11%)	6	3
All	All	445/450 (99%)	414 (93%)	31 (7%)	15	10

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	26	SER
1	D	31	LEU
1	D	56	GLN
1	D	62	LYS
1	D	87	VAL
1	D	91	LYS
1	E	3	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	14	ASN
1	E	23	LYS
1	E	35	ARG
1	E	44	ASN
1	E	56	GLN
1	E	63	LYS
1	E	82	VAL
1	F	3	GLN
1	F	31	LEU
1	G	1	THR
1	G	31	LEU
1	G	59	ASP
1	G	82	VAL
1	G	99	ILE
1	H	1	THR
1	H	3	GLN
1	H	7	ASP
1	H	20	LEU
1	H	31	LEU
1	H	44	ASN
1	H	74	ILE
1	H	77	LEU
1	H	82	VAL
1	H	103	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	16	GLN
1	E	44	ASN
1	E	94	HIS
1	E	103	ASN
1	F	18	HIS
1	F	49	GLN
1	G	4	ASN
1	H	21	ASN
1	H	44	ASN
1	H	94	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](#)

23 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	GAL	A	1	2	11,11,12	1.49	1 (9%)	15,15,17	1.69	5 (33%)
2	NGA	A	2	2	14,14,15	0.70	0	17,19,21	1.83	3 (17%)
2	GAL	A	3	2	11,11,12	1.09	0	15,15,17	1.42	1 (6%)
2	SIA	A	4	2	20,20,21	2.90	5 (25%)	24,28,31	1.73	7 (29%)
3	BGC	B	1	3	12,12,12	1.13	0	17,17,17	1.44	1 (5%)
3	GAL	B	2	3	11,11,12	2.17	4 (36%)	15,15,17	2.49	4 (26%)
3	NGA	B	3	3	14,14,15	1.62	2 (14%)	17,19,21	1.40	2 (11%)
3	GAL	B	4	3	11,11,12	1.49	2 (18%)	15,15,17	1.19	1 (6%)
3	SIA	B	5	3	20,20,21	1.68	5 (25%)	24,28,31	1.40	5 (20%)
2	GAL	C	1	2	11,11,12	1.37	2 (18%)	15,15,17	2.15	4 (26%)
2	NGA	C	2	2	14,14,15	0.91	0	17,19,21	1.71	4 (23%)
2	GAL	C	3	2	11,11,12	1.52	2 (18%)	15,15,17	1.07	1 (6%)
2	SIA	C	4	2	20,20,21	2.25	7 (35%)	24,28,31	1.56	4 (16%)
3	BGC	I	1	3	12,12,12	1.76	5 (41%)	17,17,17	2.15	6 (35%)
3	GAL	I	2	3	11,11,12	1.45	2 (18%)	15,15,17	2.52	4 (26%)
3	NGA	I	3	3	14,14,15	0.94	0	17,19,21	1.44	3 (17%)
3	GAL	I	4	3	11,11,12	1.16	1 (9%)	15,15,17	1.10	1 (6%)
3	SIA	I	5	3	20,20,21	1.86	7 (35%)	24,28,31	1.68	7 (29%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	BGC	J	1	3	12,12,12	1.47	1 (8%)	17,17,17	1.48	4 (23%)
3	GAL	J	2	3	11,11,12	0.99	0	15,15,17	1.70	3 (20%)
3	NGA	J	3	3	14,14,15	0.98	1 (7%)	17,19,21	1.33	2 (11%)
3	GAL	J	4	3	11,11,12	1.47	3 (27%)	15,15,17	0.98	1 (6%)
3	SIA	J	5	3	20,20,21	1.10	1 (5%)	24,28,31	1.47	5 (20%)

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	GAL	A	1	2	-	0/2/19/22	0/1/1/1
2	NGA	A	2	2	-	0/6/23/26	0/1/1/1
2	GAL	A	3	2	-	0/2/19/22	0/1/1/1
2	SIA	A	4	2	-	2/18/34/38	0/1/1/1
3	BGC	B	1	3	-	2/2/22/22	0/1/1/1
3	GAL	B	2	3	-	0/2/19/22	0/1/1/1
3	NGA	B	3	3	-	0/6/23/26	0/1/1/1
3	GAL	B	4	3	-	0/2/19/22	0/1/1/1
3	SIA	B	5	3	-	1/18/34/38	0/1/1/1
2	GAL	C	1	2	-	0/2/19/22	0/1/1/1
2	NGA	C	2	2	-	0/6/23/26	0/1/1/1
2	GAL	C	3	2	-	0/2/19/22	0/1/1/1
2	SIA	C	4	2	-	1/18/34/38	0/1/1/1
3	BGC	I	1	3	-	0/2/22/22	0/1/1/1
3	GAL	I	2	3	-	0/2/19/22	0/1/1/1
3	NGA	I	3	3	-	0/6/23/26	0/1/1/1
3	GAL	I	4	3	-	0/2/19/22	0/1/1/1
3	SIA	I	5	3	-	2/18/34/38	0/1/1/1
3	BGC	J	1	3	-	0/2/22/22	0/1/1/1
3	GAL	J	2	3	-	0/2/19/22	0/1/1/1
3	NGA	J	3	3	-	0/6/23/26	0/1/1/1
3	GAL	J	4	3	-	0/2/19/22	0/1/1/1
3	SIA	J	5	3	-	1/18/34/38	0/1/1/1

All (51) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	4	SIA	C2-C1	8.76	1.60	1.52

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	4	SIA	C4-C5	6.94	1.59	1.53
2	C	4	SIA	C4-C5	5.56	1.58	1.53
3	I	5	SIA	C6-C5	4.18	1.59	1.53
3	B	3	NGA	O5-C1	-4.06	1.37	1.43
2	C	4	SIA	C2-C1	3.95	1.55	1.52
3	J	1	BGC	C6-C5	3.87	1.64	1.51
3	B	2	GAL	C2-C3	3.78	1.58	1.52
3	B	2	GAL	C1-C2	3.51	1.60	1.52
2	C	4	SIA	C6-C5	3.45	1.58	1.53
3	B	5	SIA	O6-C6	3.43	1.49	1.44
3	I	2	GAL	C1-C2	3.38	1.59	1.52
3	I	1	BGC	C6-C5	3.32	1.63	1.51
2	C	3	GAL	C2-C3	3.30	1.57	1.52
3	I	1	BGC	C1-C2	3.28	1.60	1.52
3	B	3	NGA	O4-C4	-3.21	1.35	1.43
3	B	5	SIA	C2-C1	3.18	1.55	1.52
3	B	2	GAL	O5-C1	3.16	1.48	1.43
3	B	5	SIA	C9-C8	3.11	1.60	1.52
2	A	4	SIA	C3-C2	3.07	1.57	1.52
3	I	5	SIA	O6-C6	3.07	1.48	1.44
3	B	4	GAL	C4-C5	2.99	1.59	1.53
2	A	4	SIA	C3-C4	2.93	1.57	1.52
2	C	4	SIA	O8-C8	-2.90	1.37	1.43
3	J	4	GAL	O5-C1	2.86	1.48	1.43
3	B	5	SIA	C4-C5	2.84	1.55	1.53
3	I	5	SIA	C2-C1	2.80	1.54	1.52
2	C	4	SIA	C3-C2	2.78	1.57	1.52
2	C	4	SIA	O6-C2	2.76	1.47	1.43
3	I	5	SIA	C3-C2	2.72	1.57	1.52
2	A	1	GAL	C4-C5	2.69	1.58	1.53
3	I	4	GAL	O5-C1	2.59	1.47	1.43
2	C	1	GAL	O5-C5	2.59	1.48	1.43
2	A	4	SIA	C6-C5	2.50	1.57	1.53
2	C	1	GAL	O5-C1	2.48	1.47	1.43
3	J	4	GAL	C1-C2	2.45	1.57	1.52
3	J	5	SIA	C3-C2	2.43	1.56	1.52
3	I	5	SIA	O6-C2	2.43	1.47	1.43
3	I	5	SIA	C4-C5	2.41	1.55	1.53
3	B	5	SIA	C3-C2	2.39	1.56	1.52
3	J	3	NGA	C8-C7	2.37	1.55	1.50
3	I	1	BGC	O5-C1	2.28	1.48	1.42
2	C	3	GAL	C4-C5	2.26	1.57	1.53

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	2	GAL	C2-C3	2.15	1.55	1.52
3	B	4	GAL	O3-C3	-2.14	1.37	1.43
3	B	2	GAL	O5-C5	2.11	1.47	1.43
3	J	4	GAL	C4-C5	2.10	1.57	1.53
3	I	1	BGC	C4-C5	2.06	1.57	1.53
3	I	5	SIA	C3-C4	-2.06	1.49	1.52
3	I	1	BGC	O5-C5	2.06	1.49	1.44
2	C	4	SIA	C11-C10	2.04	1.54	1.50

All (78) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	2	GAL	C1-O5-C5	7.41	122.24	112.19
3	I	2	GAL	C1-O5-C5	6.83	121.44	112.19
2	A	2	NGA	C1-O5-C5	5.04	119.02	112.19
2	C	1	GAL	C1-O5-C5	4.80	118.70	112.19
3	I	1	BGC	C3-C4-C5	4.51	118.29	110.24
3	I	5	SIA	C4-C5-N5	-4.31	101.83	110.38
3	B	1	BGC	C1-O5-C5	4.24	121.66	113.66
3	J	2	GAL	O2-C2-C1	4.23	117.81	109.15
3	B	2	GAL	C2-C3-C4	4.19	118.15	110.89
2	C	1	GAL	O3-C3-C2	-4.12	102.11	109.99
3	I	2	GAL	O5-C5-C6	-4.05	100.86	107.20
2	C	2	NGA	C1-O5-C5	3.97	117.57	112.19
3	I	1	BGC	O2-C2-C1	3.95	118.31	109.16
2	A	2	NGA	C1-C2-N2	-3.88	103.86	110.49
2	C	1	GAL	O2-C2-C1	3.82	116.97	109.15
3	J	3	NGA	C1-C2-N2	-3.82	103.96	110.49
2	A	4	SIA	O6-C2-C3	-3.61	105.48	110.46
3	J	2	GAL	O3-C3-C2	-3.61	103.08	109.99
2	A	3	GAL	C1-O5-C5	3.55	117.01	112.19
3	I	2	GAL	O3-C3-C2	-3.51	103.28	109.99
2	A	4	SIA	C9-C8-C7	-3.43	104.97	112.41
3	B	4	GAL	C1-O5-C5	3.39	116.78	112.19
3	J	5	SIA	C4-C5-N5	-3.37	103.71	110.38
2	C	4	SIA	O6-C2-C1	3.37	114.30	107.70
3	I	4	GAL	C1-C2-C3	-3.14	105.81	109.67
3	B	3	NGA	C1-C2-N2	-3.13	105.15	110.49
3	B	2	GAL	O3-C3-C4	-3.12	103.14	110.35
3	I	2	GAL	O5-C1-C2	3.06	115.49	110.77
3	J	1	BGC	O4-C4-C3	-3.00	103.40	110.35
2	C	2	NGA	O7-C7-C8	-2.88	116.72	122.06

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	5	SIA	O1B-C1-C2	2.84	121.15	113.03
2	A	1	GAL	O4-C4-C3	-2.83	103.81	110.35
3	J	1	BGC	O5-C5-C4	-2.79	104.64	109.69
2	C	4	SIA	O1B-C1-C2	2.75	120.87	113.03
2	A	1	GAL	C2-C3-C4	-2.73	106.17	110.89
2	C	3	GAL	C6-C5-C4	2.73	119.40	113.00
2	A	4	SIA	C6-O6-C2	2.72	117.16	111.34
3	I	5	SIA	C3-C4-C5	-2.70	108.20	111.46
2	A	1	GAL	C1-C2-C3	2.64	112.91	109.67
3	I	1	BGC	O5-C5-C6	2.63	112.98	106.44
3	I	1	BGC	O4-C4-C5	-2.62	102.79	109.30
2	C	2	NGA	C8-C7-N2	2.62	120.53	116.10
3	B	5	SIA	C4-C5-N5	-2.56	105.30	110.38
3	B	5	SIA	O1A-C1-C2	-2.56	116.51	122.57
3	B	2	GAL	O3-C3-C2	-2.55	105.11	109.99
3	J	5	SIA	C5-N5-C10	-2.54	117.01	123.18
2	A	4	SIA	O8-C8-C7	2.51	115.21	109.10
3	I	5	SIA	O1B-C1-C2	2.49	120.13	113.03
2	A	2	NGA	C2-N2-C7	2.46	126.41	122.90
3	B	5	SIA	C6-O6-C2	2.42	116.52	111.34
2	A	4	SIA	O1B-C1-O1A	-2.41	118.61	124.09
3	B	5	SIA	C3-C4-C5	-2.41	108.55	111.46
3	J	1	BGC	O5-C1-C2	-2.41	105.99	110.28
2	C	4	SIA	C4-C5-N5	-2.39	105.64	110.38
3	I	5	SIA	O1A-C1-C2	-2.38	116.96	122.57
3	I	3	NGA	C1-C2-N2	-2.32	106.53	110.49
2	A	4	SIA	O1B-C1-C2	2.32	119.64	113.03
2	A	4	SIA	C3-C4-C5	-2.31	108.67	111.46
3	J	5	SIA	O1B-C1-C2	2.30	119.60	113.03
3	I	1	BGC	O3-C3-C2	2.25	115.54	110.35
3	J	1	BGC	O5-C5-C6	2.24	112.02	106.44
3	I	3	NGA	O6-C6-C5	-2.23	103.63	111.29
3	B	3	NGA	C6-C5-C4	2.23	118.23	113.00
3	I	5	SIA	O8-C8-C7	2.22	114.50	109.10
3	I	1	BGC	O1-C1-C2	2.20	115.22	109.03
3	J	5	SIA	C3-C4-C5	-2.19	108.81	111.46
3	I	3	NGA	O5-C5-C6	2.17	110.61	107.20
3	J	2	GAL	C1-O5-C5	2.17	115.13	112.19
2	A	1	GAL	O3-C3-C2	-2.13	105.91	109.99
2	C	1	GAL	O4-C4-C3	-2.12	105.44	110.35
2	C	4	SIA	O1A-C1-C2	-2.12	117.57	122.57
2	A	1	GAL	C6-C5-C4	2.10	117.91	113.00

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	5	SIA	O6-C2-C1	2.08	111.78	107.70
3	I	5	SIA	O6-C2-C3	-2.08	107.60	110.46
3	J	3	NGA	C1-O5-C5	2.08	115.00	112.19
3	J	4	GAL	O2-C2-C1	2.05	113.35	109.15
2	C	2	NGA	C4-C3-C2	2.02	113.98	111.02
3	J	5	SIA	C6-O6-C2	2.02	115.65	111.34

There are no chirality outliers.

All (9) torsion outliers are listed below:

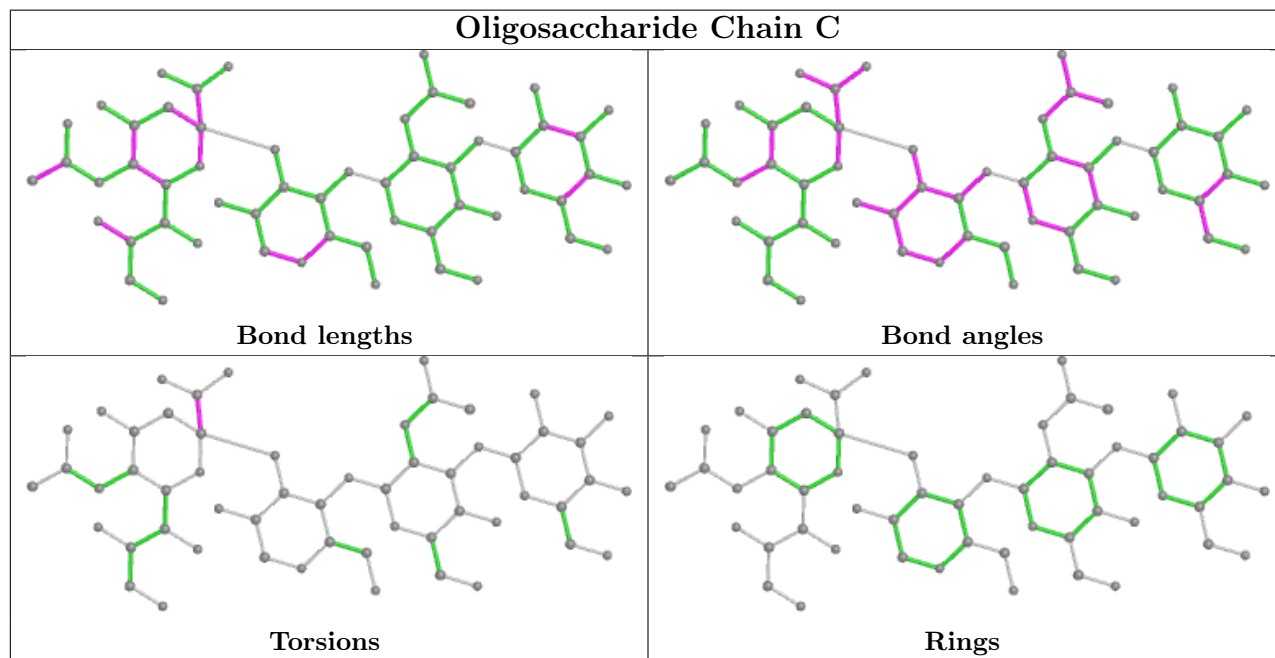
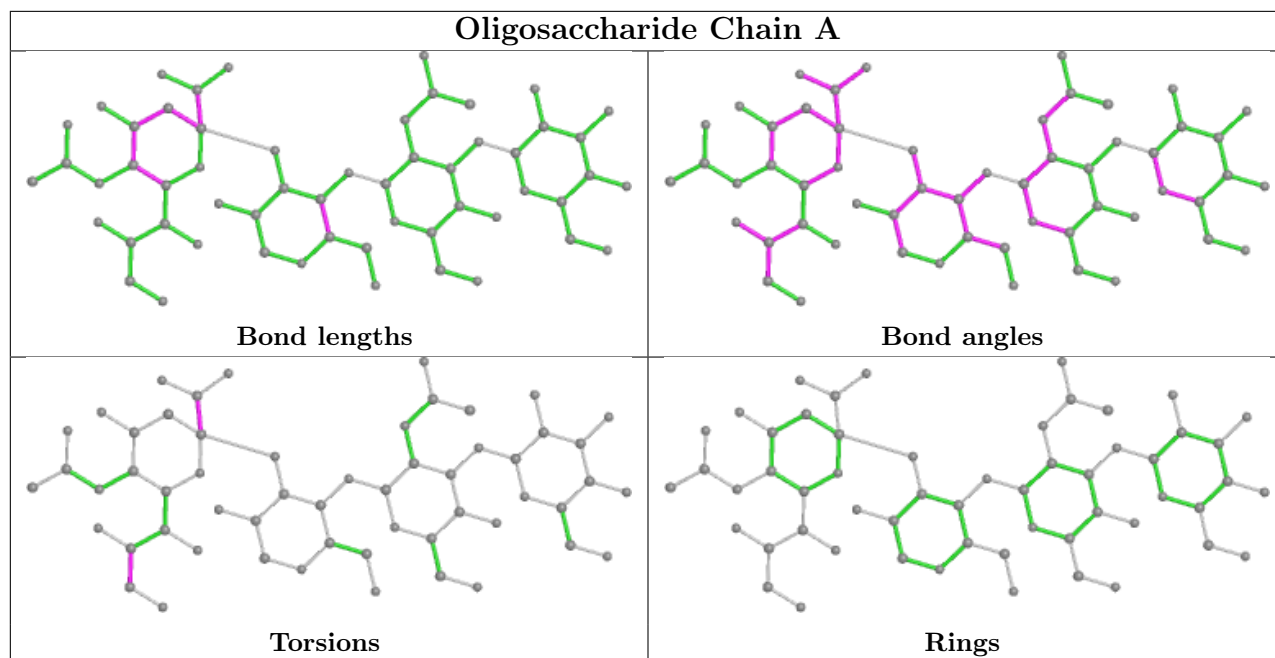
Mol	Chain	Res	Type	Atoms
3	B	1	BGC	O5-C5-C6-O6
3	B	1	BGC	C4-C5-C6-O6
2	A	4	SIA	O8-C8-C9-O9
2	A	4	SIA	O1A-C1-C2-O6
2	C	4	SIA	O1A-C1-C2-O6
3	B	5	SIA	O1A-C1-C2-O6
3	I	5	SIA	O1A-C1-C2-O6
3	J	5	SIA	O1A-C1-C2-O6
3	I	5	SIA	O8-C8-C9-O9

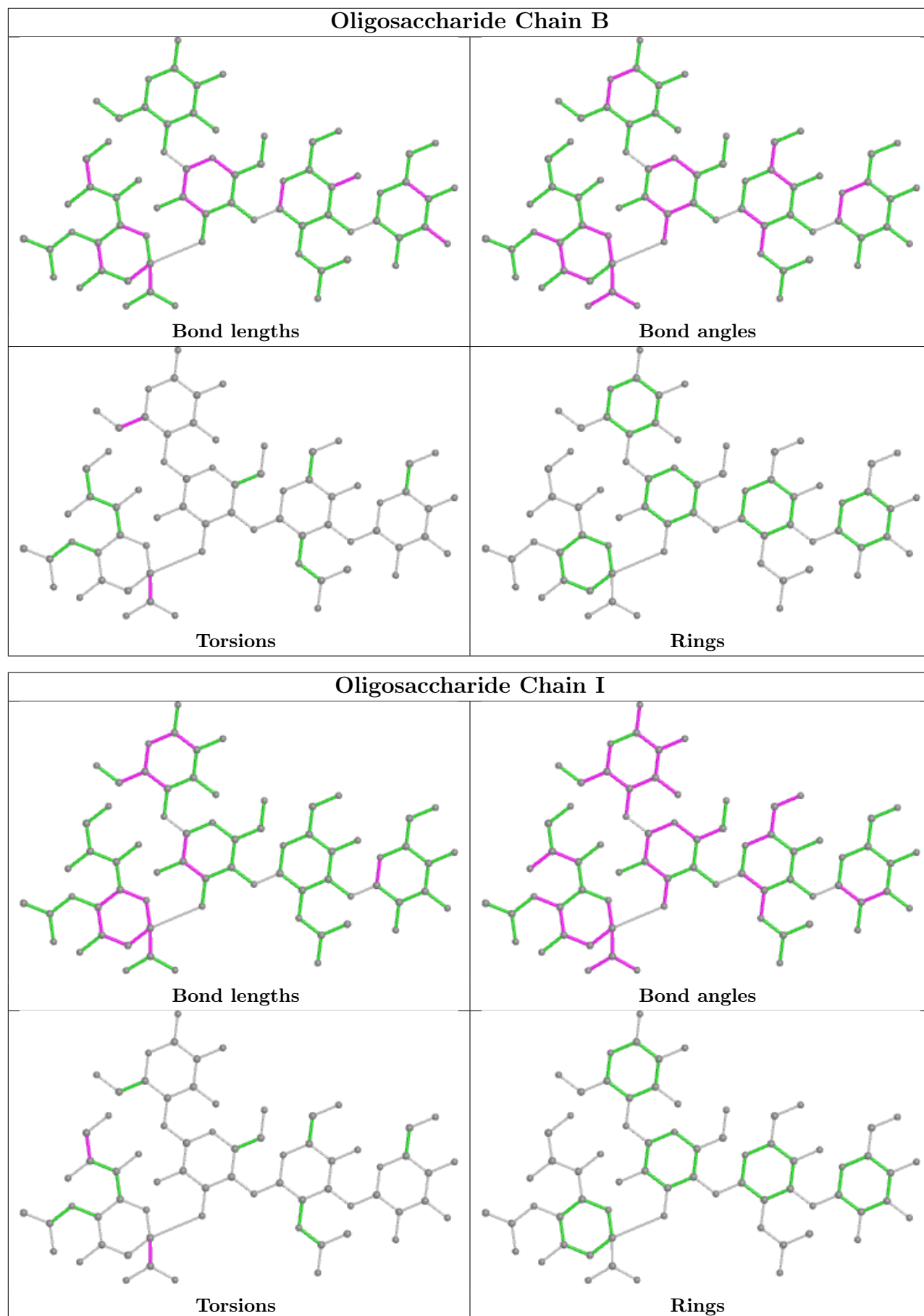
There are no ring outliers.

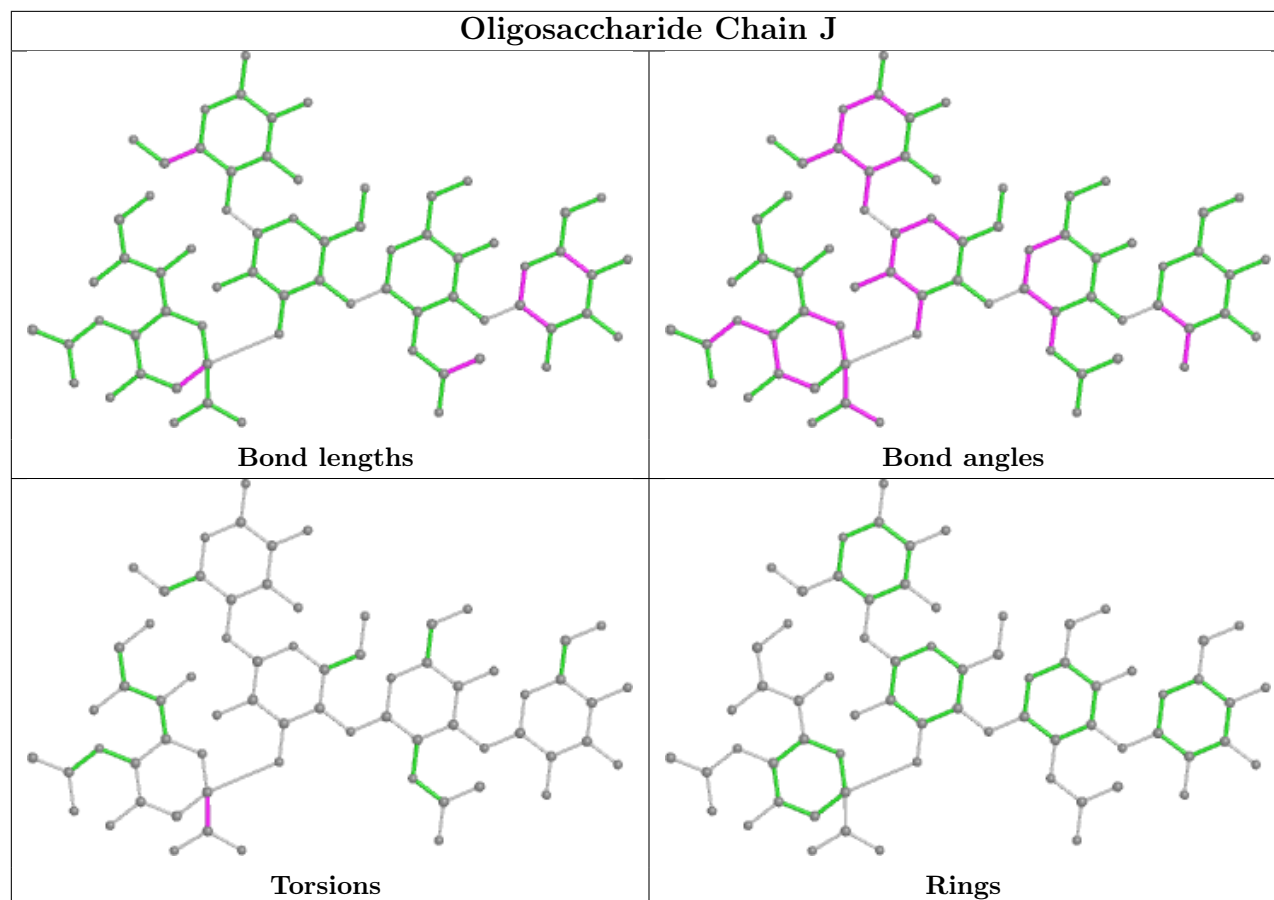
2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	J	4	GAL	1	0
2	A	4	SIA	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](#)

There are no ligands in this entry.

## 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	D	103/104 (99%)	-0.20	0 <b>100</b> <b>100</b>	9, 22, 43, 53	1 (0%)
1	E	103/104 (99%)	-0.32	1 (0%) 82 81	9, 20, 36, 60	0
1	F	103/104 (99%)	-0.20	2 (1%) 66 65	9, 26, 45, 56	0
1	G	103/104 (99%)	-0.47	1 (0%) 82 81	8, 17, 32, 55	0
1	H	103/104 (99%)	-0.46	0 <b>100</b> <b>100</b>	7, 16, 32, 56	0
All	All	515/520 (99%)	-0.33	4 (0%) <b>86</b> <b>85</b>	7, 19, 41, 60	1 (0%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	103	ASN	2.4
1	E	103	ASN	2.3
1	F	103	ASN	2.3
1	F	15	THR	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	GAL	A	1	11/12	0.60	0.20	35,42,46,48	11

*Continued on next page...*



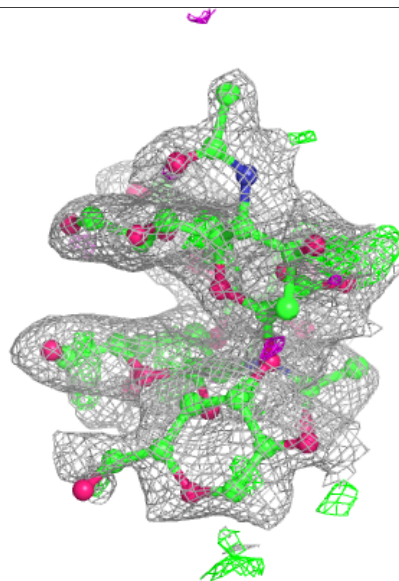
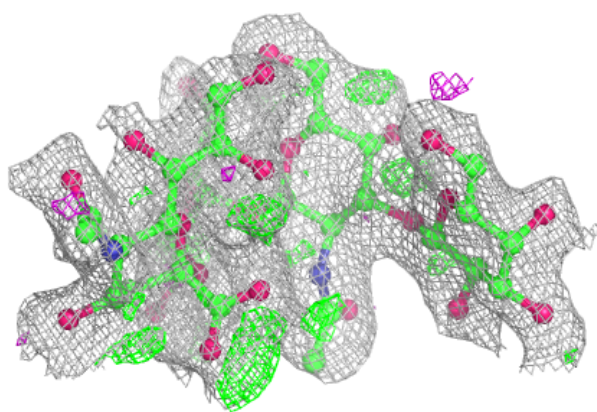
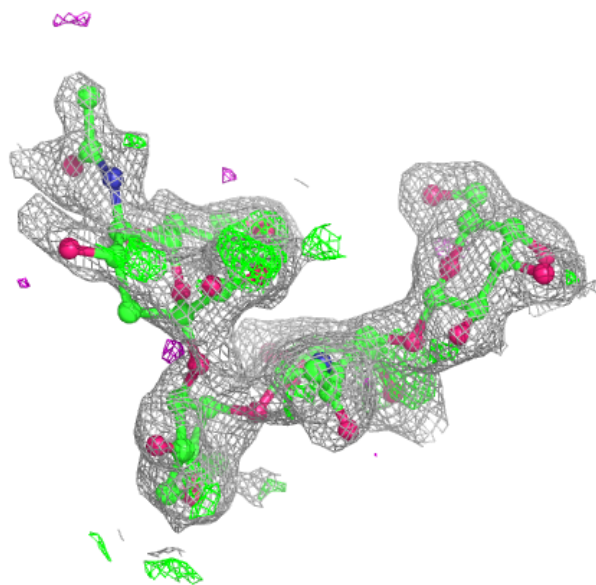
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	GAL	C	1	11/12	0.68	0.13	41,45,50,52	0
2	SIA	A	4	20/21	0.75	0.17	31,37,42,44	20
2	NGA	A	2	14/15	0.75	0.16	23,25,27,33	14
3	BGC	I	1	12/12	0.77	0.14	43,45,48,48	0
3	BGC	B	1	12/12	0.78	0.17	45,52,56,59	0
2	SIA	C	4	20/21	0.81	0.14	29,35,39,42	0
2	GAL	C	3	11/12	0.83	0.11	23,34,37,37	0
3	SIA	B	5	20/21	0.84	0.13	23,30,34,36	0
2	NGA	C	2	14/15	0.84	0.15	34,39,40,42	0
3	GAL	I	2	11/12	0.84	0.13	30,39,43,48	0
3	BGC	J	1	12/12	0.85	0.14	29,40,43,44	0
3	SIA	I	5	20/21	0.86	0.10	20,25,28,28	0
2	GAL	A	3	11/12	0.86	0.11	18,21,23,24	11
3	GAL	J	2	11/12	0.86	0.13	20,23,27,33	0
3	GAL	B	2	11/12	0.88	0.11	30,36,42,44	0
3	NGA	I	3	14/15	0.89	0.09	22,25,30,33	0
3	NGA	B	3	14/15	0.90	0.10	18,23,26,27	0
3	SIA	J	5	20/21	0.91	0.09	13,20,25,27	0
3	GAL	B	4	11/12	0.92	0.10	11,18,23,24	0
3	GAL	J	4	11/12	0.92	0.08	12,15,18,19	0
3	GAL	I	4	11/12	0.92	0.08	12,15,19,20	0
3	NGA	J	3	14/15	0.95	0.08	16,17,22,24	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.

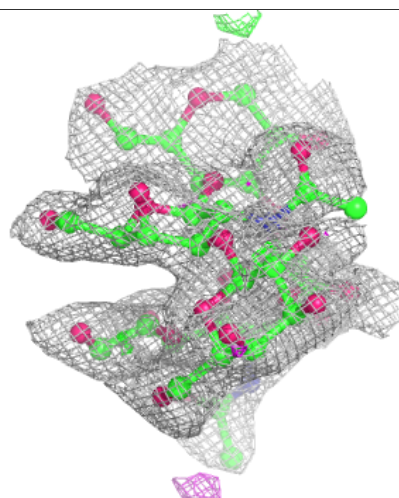
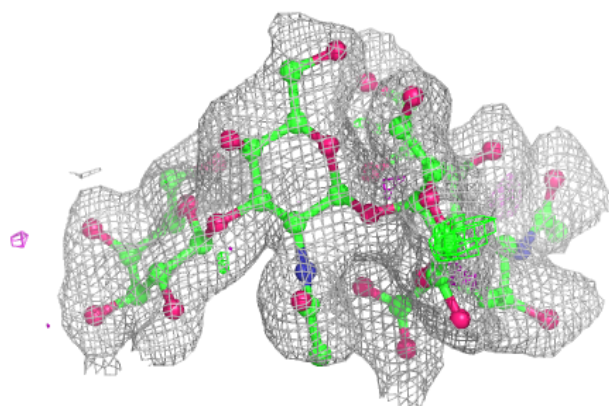
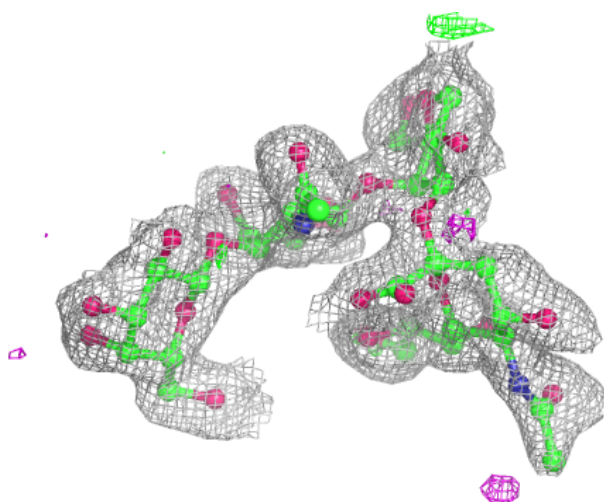
**Electron density around Chain A:**

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



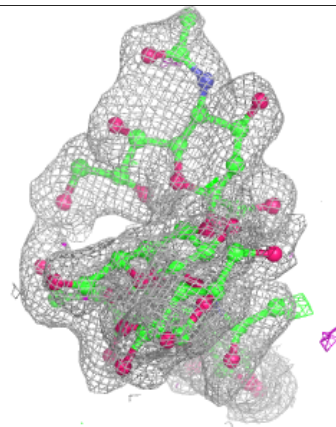
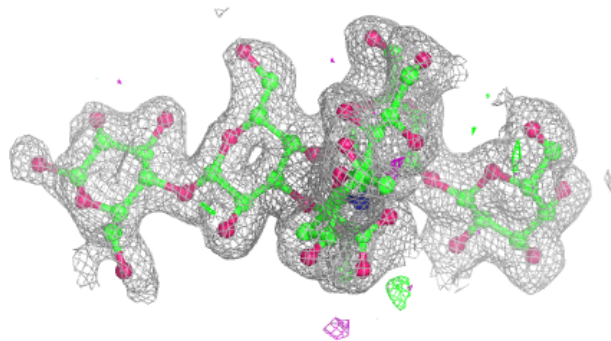
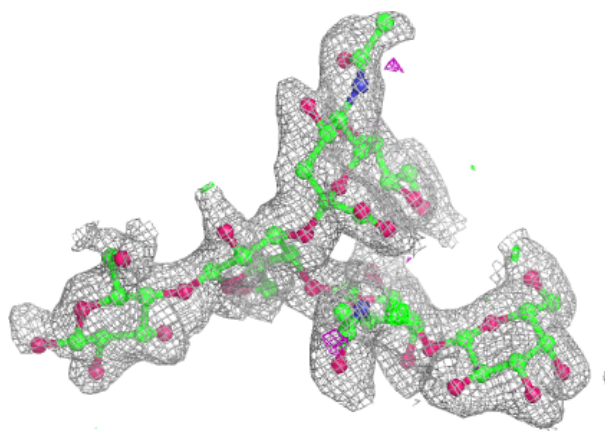
**Electron density around Chain C:**

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



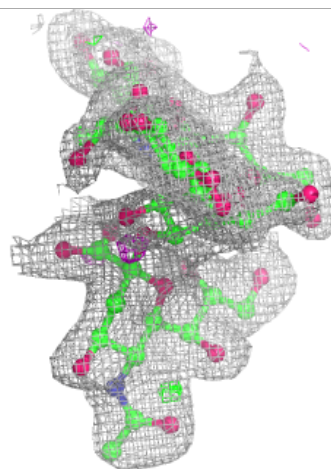
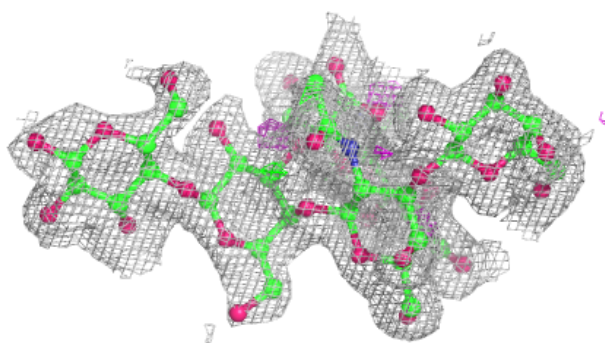
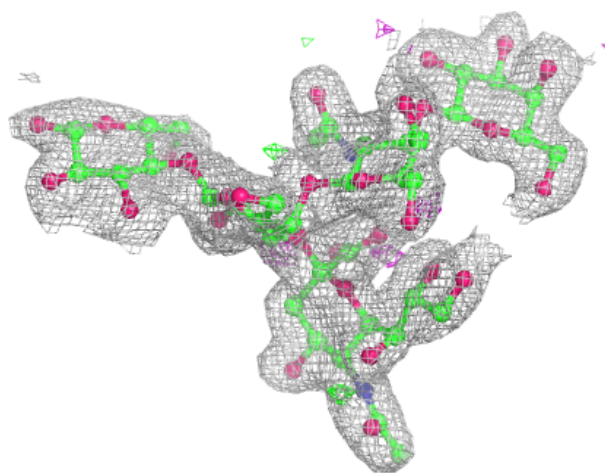
**Electron density around Chain B:**

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

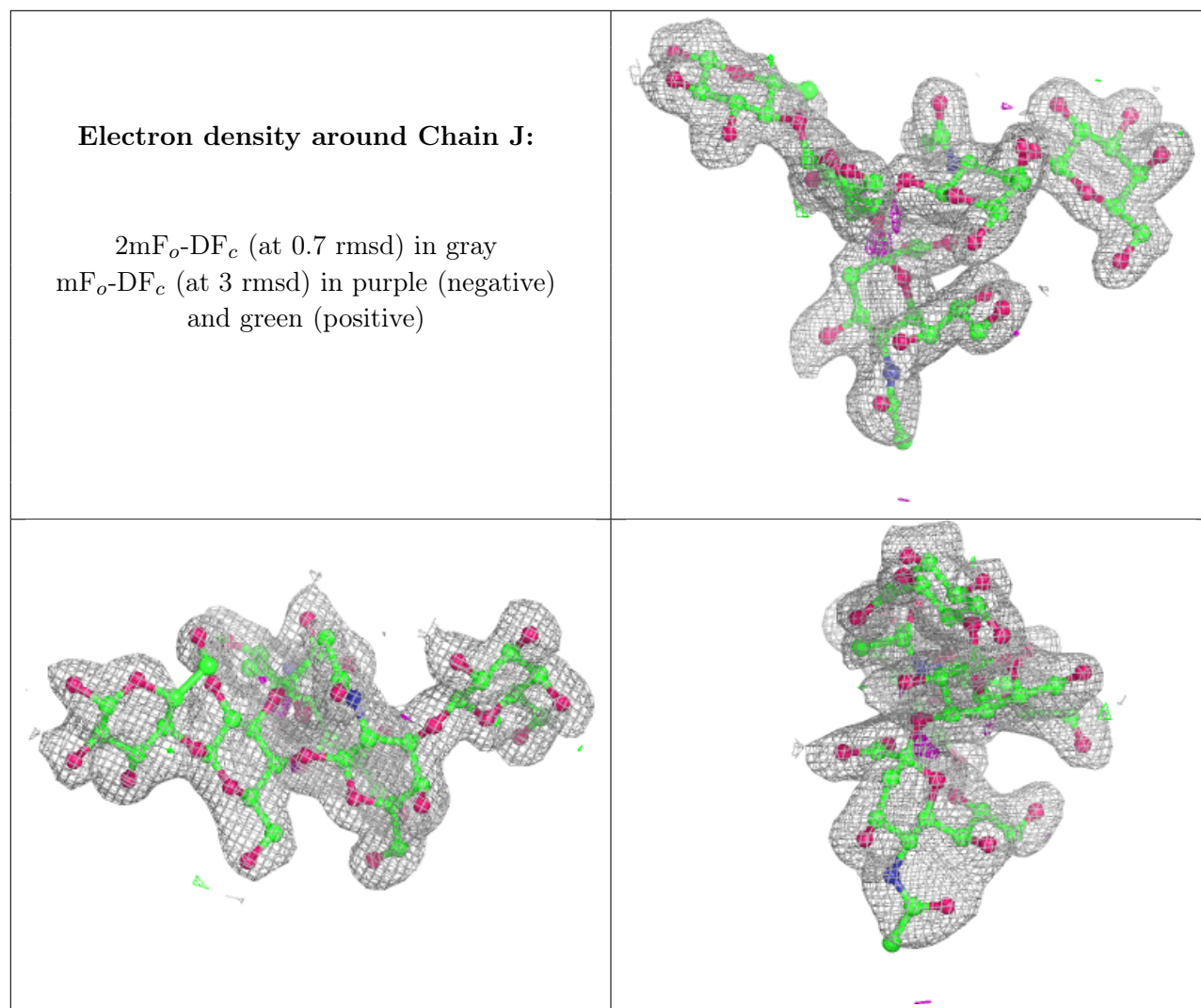


**Electron density around Chain I:**

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







## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.