



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 30, 2022 – 02:12 PM JST

PDB ID : 7VN5  
Title : Crystal structure of MBP-fused BIL1/BZR1 (21-90) in complex with double-stranded DNA containing TTCACGTGAA  
Authors : Nosaki, S.; Tanokura, M.; Miyakawa, T.  
Deposited on : 2021-10-10  
Resolution : 1.95 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

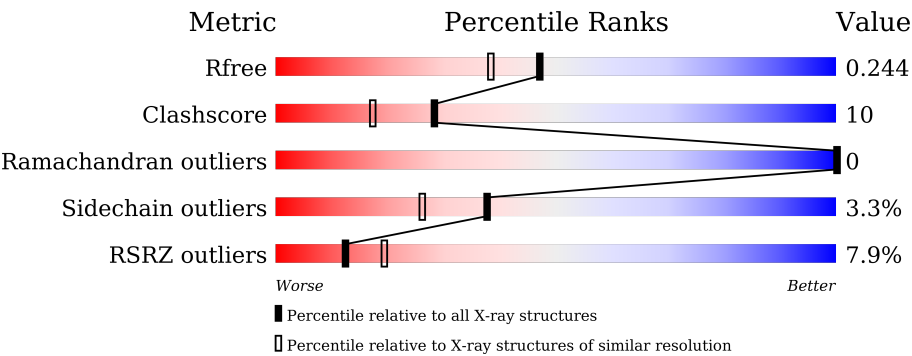
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.3
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

# 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 1.95 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.



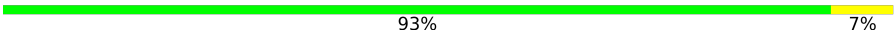



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	439	<div><div>12%</div><div>80%</div><div>18%</div><div>..</div></div>
1	B	439	<div><div>18%</div><div>75%</div><div>22%</div><div>..</div></div>
1	C	439	<div><div>%</div><div>91%</div><div>8%</div><div>.</div></div>
1	D	439	<div><div>2%</div><div>88%</div><div>10%</div><div>.</div></div>
2	E	15	<div><div>7%</div><div>60%</div><div>33%</div><div>7%</div></div>
2	F	15	<div><div>40%</div><div>60%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	15	 93% 7%
2	H	15	 87% 13%
3	K	2	 50% 50%
3	L	2	 50% 50%
3	M	2	 100%
3	N	2	 100%

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
4	EDO	D	102	-	-	X	-
4	EDO	D	104	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15403 atoms, of which 36 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 Maltodextrin-binding protein, Protein BRASSINAZOLE-RESISTANT 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	437	Total	C	N	O	S	0	0	0
			3402	2171	582	640	9			
1	D	437	Total	C	N	O	S	0	0	0
			3402	2171	582	640	9			
1	A	436	Total	C	N	O	S	0	0	0
			3394	2166	581	639	8			
1	B	436	Total	C	N	O	S	0	0	0
			3394	2166	581	639	8			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-348	MET	-	initiating methionine	UNP A0A4P1LXE0
C	-266	ALA	ASP	engineered mutation	UNP A0A4P1LXE0
C	-265	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
C	-176	ALA	GLU	engineered mutation	UNP A0A4P1LXE0
C	-175	ALA	ASN	engineered mutation	UNP A0A4P1LXE0
C	-109	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
C	11	ALA	GLU	engineered mutation	UNP A0A4P1LXE0
C	14	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
C	15	ALA	ASP	engineered mutation	UNP A0A4P1LXE0
D	-348	MET	-	initiating methionine	UNP A0A4P1LXE0
D	-266	ALA	ASP	engineered mutation	UNP A0A4P1LXE0
D	-265	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
D	-176	ALA	GLU	engineered mutation	UNP A0A4P1LXE0
D	-175	ALA	ASN	engineered mutation	UNP A0A4P1LXE0
D	-109	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
D	11	ALA	GLU	engineered mutation	UNP A0A4P1LXE0
D	14	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
D	15	ALA	ASP	engineered mutation	UNP A0A4P1LXE0
A	-348	MET	-	initiating methionine	UNP A0A4P1LXE0
A	-266	ALA	ASP	engineered mutation	UNP A0A4P1LXE0

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-265	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
A	-176	ALA	GLU	engineered mutation	UNP A0A4P1LXE0
A	-175	ALA	ASN	engineered mutation	UNP A0A4P1LXE0
A	-109	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
A	11	ALA	GLU	engineered mutation	UNP A0A4P1LXE0
A	14	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
A	15	ALA	ASP	engineered mutation	UNP A0A4P1LXE0
B	-348	MET	-	initiating methionine	UNP A0A4P1LXE0
B	-266	ALA	ASP	engineered mutation	UNP A0A4P1LXE0
B	-265	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
B	-176	ALA	GLU	engineered mutation	UNP A0A4P1LXE0
B	-175	ALA	ASN	engineered mutation	UNP A0A4P1LXE0
B	-109	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
B	11	ALA	GLU	engineered mutation	UNP A0A4P1LXE0
B	14	ALA	LYS	engineered mutation	UNP A0A4P1LXE0
B	15	ALA	ASP	engineered mutation	UNP A0A4P1LXE0

- Molecule 2 is a DNA chain called DNA (5'-D(\*TP\*TP\*TP\*TP\*CP\*AP\*CP\*GP\*TP\*GP\*AP\*AP\*AP\*AP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	15	Total	C	N	O	P	0	0	0
			305	148	56	87	14			
2	H	15	Total	C	N	O	P	0	0	0
			305	148	56	87	14			
2	E	15	Total	C	N	O	P	0	0	0
			305	148	56	87	14			
2	F	15	Total	C	N	O	P	0	0	0
			305	148	56	87	14			

- Molecule 3 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.



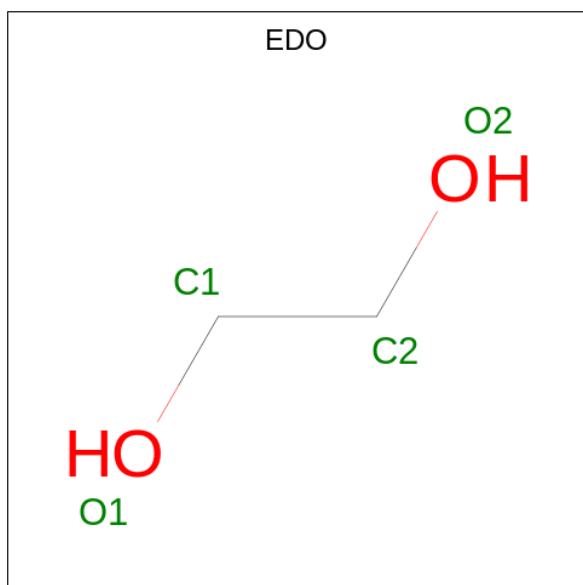
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
3	K	2	Total	C	O	0	0	0
			23	12	11			
3	L	2	Total	C	O	0	0	0
			23	12	11			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
3	M	2	Total	C	O	0	0	0
			23	12	11			
3	N	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	C	1	Total	C	O		0	0
			4	2	2			
4	C	1	Total	C	O		0	0
			4	2	2			
4	C	1	Total	C	O		0	0
			4	2	2			
4	C	1	Total	C	H	O	0	0
			10	2	6	2		
4	D	1	Total	C	O		0	0
			4	2	2			
4	D	1	Total	C	H	O	0	0
			10	2	6	2		
4	D	1	Total	C	H	O	0	0
			10	2	6	2		
4	D	1	Total	C	H	O	0	0
			10	2	6	2		
4	A	1	Total	C	O		0	0
			4	2	2			

Continued on next page...

*Continued from previous page...*

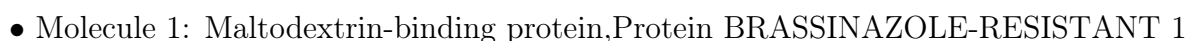
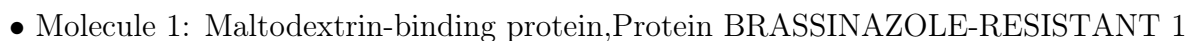
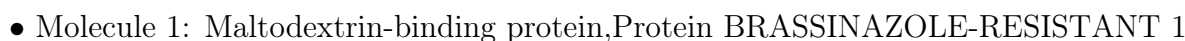
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C H O 10 2 6 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	B	1	Total C H O 10 2 6 2	0	0

- Molecule 5 is water.

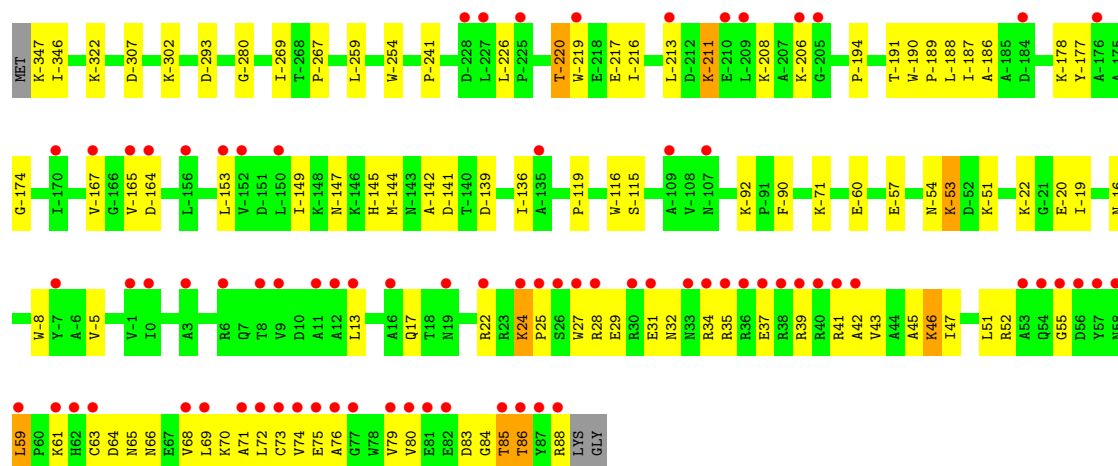
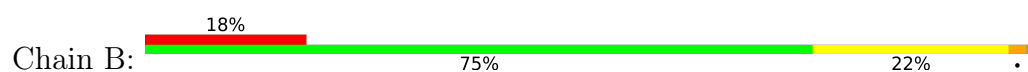
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	109	Total O 109 109	0	0
5	D	107	Total O 107 107	0	0
5	G	26	Total O 26 26	0	0
5	H	21	Total O 21 21	0	0
5	A	81	Total O 81 81	0	0
5	B	64	Total O 64 64	0	0
5	E	2	Total O 2 2	0	0
5	F	1	Total O 1 1	0	0



- Molecule 1: Maltodextrin-binding protein,Protein BRASSINAZOLE-RESISTANT 1







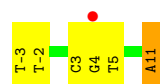
- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*TP\*CP\*AP\*CP\*GP\*TP\*GP\*AP\*AP\*AP\*AP\*A)-3')



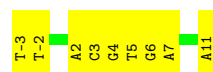
- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*TP\*CP\*AP\*CP\*GP\*TP\*GP\*AP\*AP\*AP\*AP\*A)-3')



- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*TP\*CP\*AP\*CP\*GP\*TP\*GP\*AP\*AP\*AP\*AP\*A)-3')



- Molecule 2: DNA (5'-D(\*TP\*TP\*TP\*TP\*CP\*AP\*CP\*GP\*TP\*GP\*AP\*AP\*AP\*AP\*A)-3')



- Molecule 3: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain K:  50% 50%

GLC1  
GLC2

- Molecule 3: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain L:  50% 50%

GLC1  
GLC2

- Molecule 3: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain M:  100%

GLC1  
GLC2

- Molecule 3: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain N:  100%

GLC1  
GLC2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.62Å 93.92Å 112.75Å 90.00° 100.65° 90.00°	Depositor
Resolution (Å)	35.82 – 1.95 35.82 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.2 (35.82-1.95) 99.2 (35.82-1.95)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.72 (at 1.95Å)	Xtriage
Refinement program	PHENIX 1.18_3855	Depositor
R, $R_{free}$	0.209 , 0.244 0.209 , 0.244	Depositor DCC
$R_{free}$ test set	7634 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.1	Xtriage
Anisotropy	0.481	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 39.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	15403	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 81.65 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.3753e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GLC, EDO

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	0.46	0/3473	0.51	0/4713
1	B	0.52	0/3473	0.55	0/4713
1	C	0.36	0/3481	0.47	0/4723
1	D	0.34	0/3481	0.48	0/4723
2	E	0.67	0/342	1.02	1/526 (0.2%)
2	F	0.57	0/342	1.02	1/526 (0.2%)
2	G	0.85	0/342	1.07	1/526 (0.2%)
2	H	0.81	0/342	1.09	1/526 (0.2%)
All	All	0.46	0/15276	0.58	4/20976 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	11	DA	C1'-O4'-C4'	-6.28	103.82	110.10
2	E	11	DA	C1'-O4'-C4'	-5.83	104.27	110.10
2	G	11	DA	C1'-O4'-C4'	-5.66	104.44	110.10
2	F	11	DA	C1'-O4'-C4'	-5.25	104.84	110.10

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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	3394	0	3365	113	0
1	B	3394	0	3365	124	0
1	C	3402	0	3374	33	0
1	D	3402	0	3374	47	0
2	E	305	0	172	17	0
2	F	305	0	172	10	0
2	G	305	0	172	0	0
2	H	305	0	172	1	0
3	K	23	0	21	0	0
3	L	23	0	21	0	0
3	M	23	0	21	0	0
3	N	23	0	21	0	0
4	A	12	6	18	4	0
4	B	8	6	12	2	0
4	C	16	6	24	4	0
4	D	16	18	24	9	0
5	A	81	0	0	1	0
5	B	64	0	0	3	0
5	C	109	0	0	2	0
5	D	107	0	0	3	0
5	E	2	0	0	0	0
5	F	1	0	0	0	0
5	G	26	0	0	0	0
5	H	21	0	0	1	0
All	All	15367	36	14328	284	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 10.

The worst 5 of 284 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:85:THR:HG22	1:B:46:LYS:CE	1.57	1.35
1:A:85:THR:CG2	1:B:46:LYS:HE2	1.66	1.23
1:A:69:LEU:HD21	1:B:47:ILE:HG21	1.25	1.15
1:C:-346:ILE:HD13	1:C:-77:PRO:HD3	1.21	1.12
1:A:50:GLY:HA3	1:B:86:THR:CG2	1.80	1.12

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	434/439 (99%)	426 (98%)	8 (2%)	0	100	100
1	B	434/439 (99%)	428 (99%)	6 (1%)	0	100	100
1	C	435/439 (99%)	431 (99%)	4 (1%)	0	100	100
1	D	435/439 (99%)	428 (98%)	7 (2%)	0	100	100
All	All	1738/1756 (99%)	1713 (99%)	25 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	343/345 (99%)	329 (96%)	14 (4%)	30	18
1	B	343/345 (99%)	328 (96%)	15 (4%)	28	15
1	C	344/345 (100%)	337 (98%)	7 (2%)	55	48
1	D	344/345 (100%)	335 (97%)	9 (3%)	46	36
All	All	1374/1380 (100%)	1329 (97%)	45 (3%)	38	26

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	LEU
1	B	-53	LYS
1	A	86	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	-220	THR
1	B	22	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	17	GLN
1	A	32	ASN
1	A	58	ASN
1	A	65	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

8 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	GLC	K	1	3	12,12,12	0.46	0	17,17,17	0.79	0
3	GLC	K	2	3	11,11,12	0.59	0	15,15,17	0.86	1 (6%)
3	GLC	L	1	3	12,12,12	0.53	0	17,17,17	0.99	1 (5%)
3	GLC	L	2	3	11,11,12	0.74	0	15,15,17	0.91	0
3	GLC	M	1	3	12,12,12	0.48	0	17,17,17	1.03	1 (5%)
3	GLC	M	2	3	11,11,12	0.60	0	15,15,17	1.00	2 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GLC	N	1	3	12,12,12	0.55	0	17,17,17	0.80	1 (5%)
3	GLC	N	2	3	11,11,12	0.65	0	15,15,17	0.84	1 (6%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GLC	K	1	3	-	0/2/22/22	0/1/1/1
3	GLC	K	2	3	-	0/2/19/22	0/1/1/1
3	GLC	L	1	3	-	0/2/22/22	0/1/1/1
3	GLC	L	2	3	-	0/2/19/22	0/1/1/1
3	GLC	M	1	3	-	0/2/22/22	0/1/1/1
3	GLC	M	2	3	-	0/2/19/22	0/1/1/1
3	GLC	N	1	3	-	0/2/22/22	0/1/1/1
3	GLC	N	2	3	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	M	1	GLC	O5-C1-C2	2.92	115.49	110.28
3	L	1	GLC	O5-C1-C2	2.58	114.89	110.28
3	K	2	GLC	C1-O5-C5	2.30	115.31	112.19
3	M	2	GLC	C1-O5-C5	2.29	115.29	112.19
3	M	2	GLC	O5-C5-C6	2.18	110.62	107.20

There are no chirality outliers.

There are no torsion outliers.

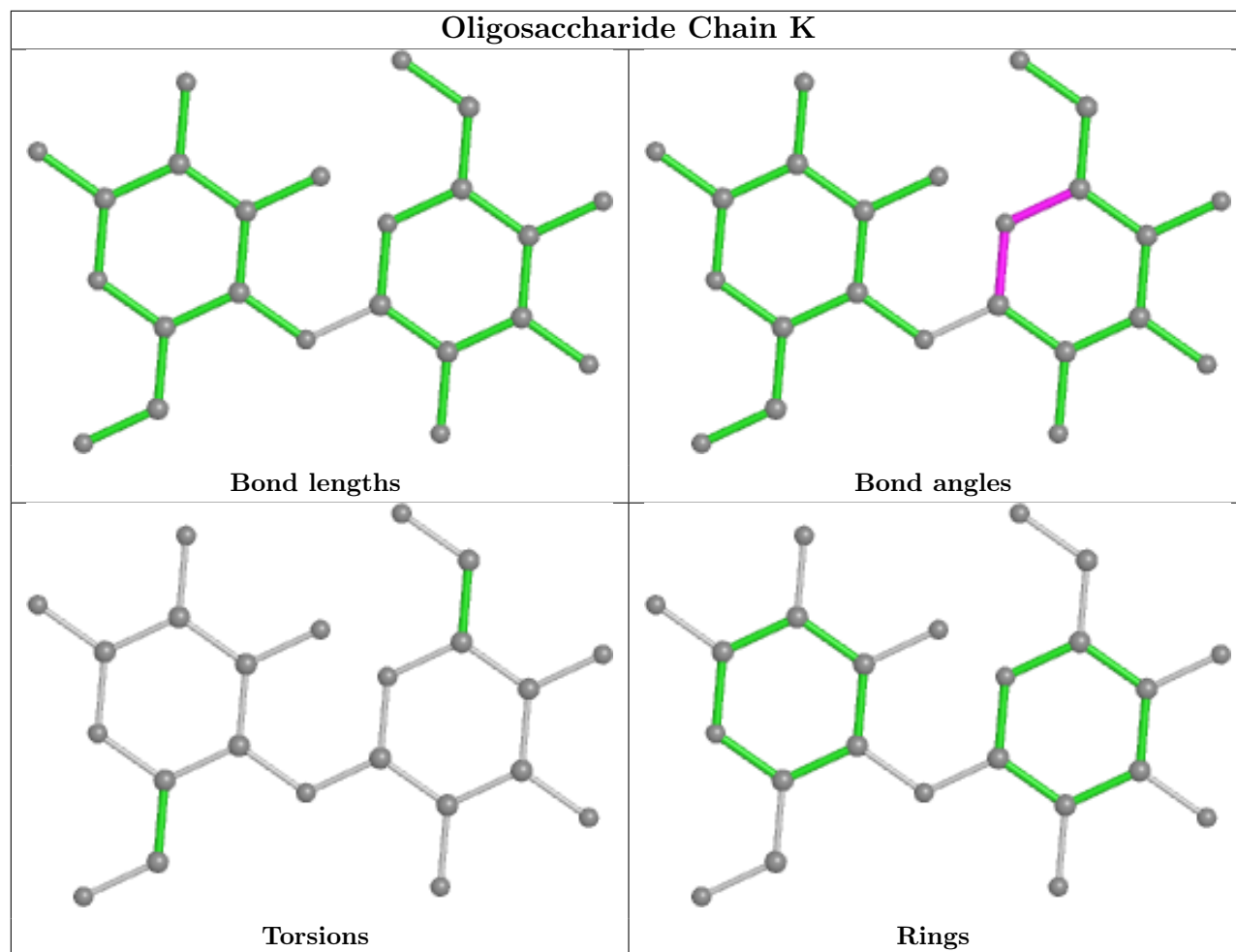
There are no ring outliers.

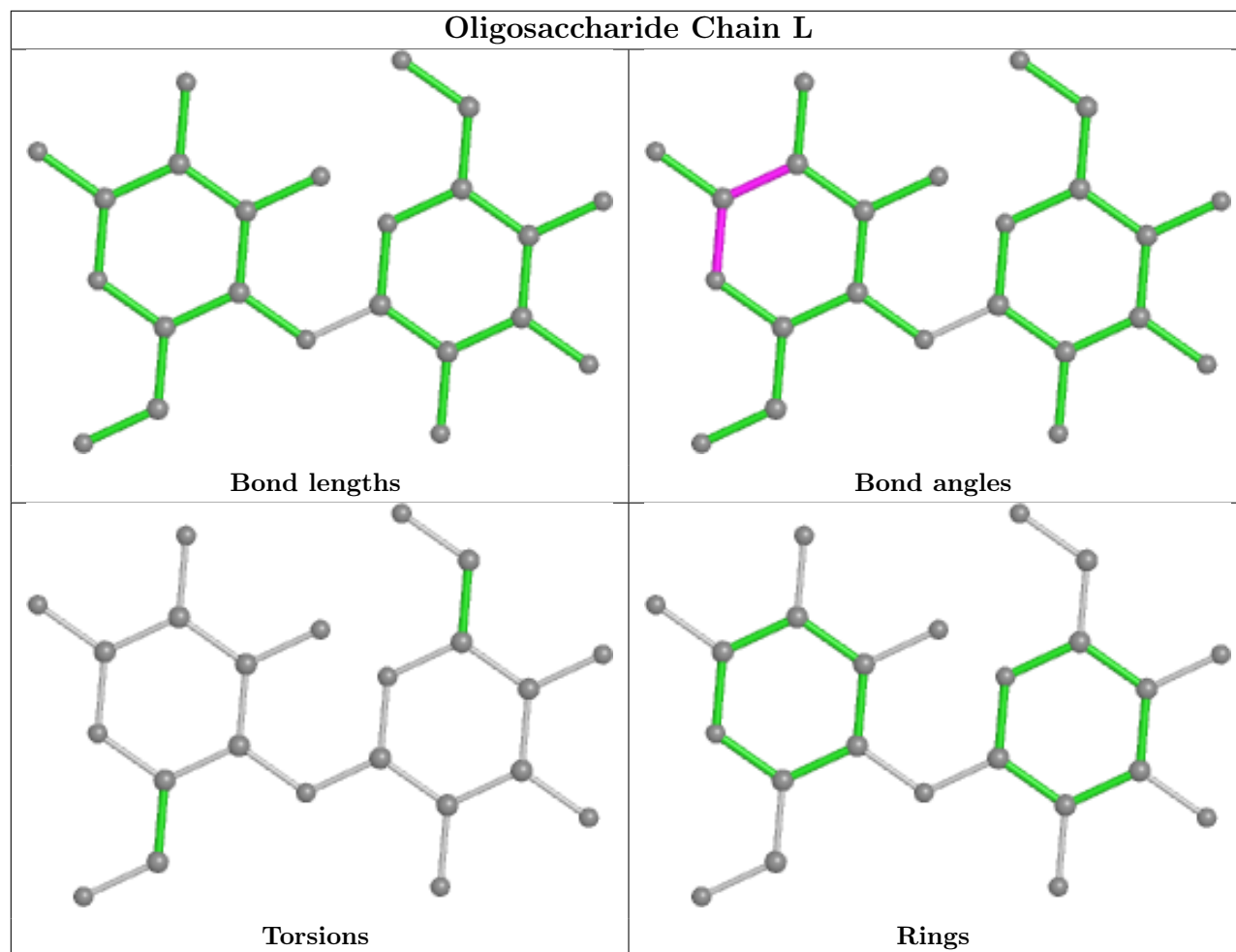
No monomer is involved in short contacts.

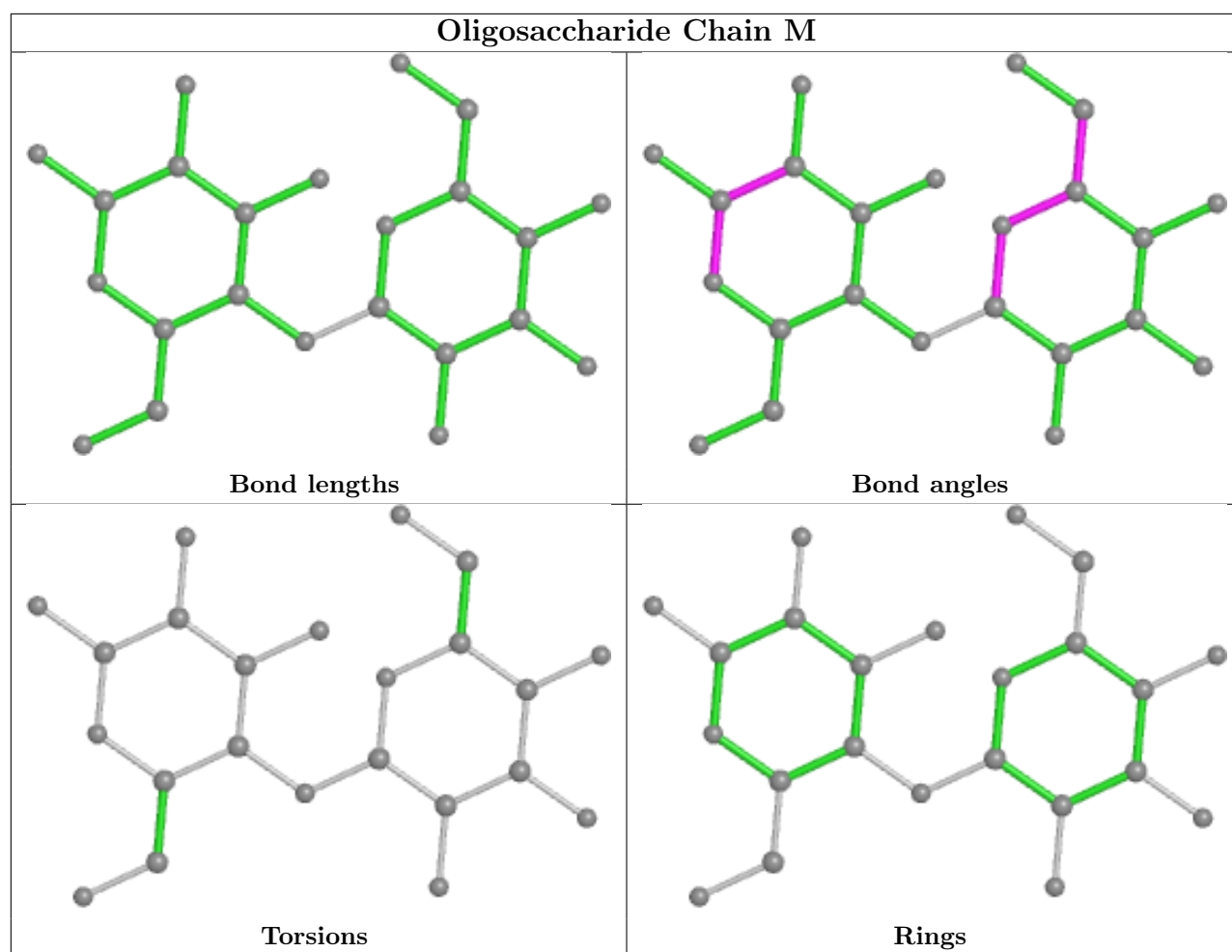
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

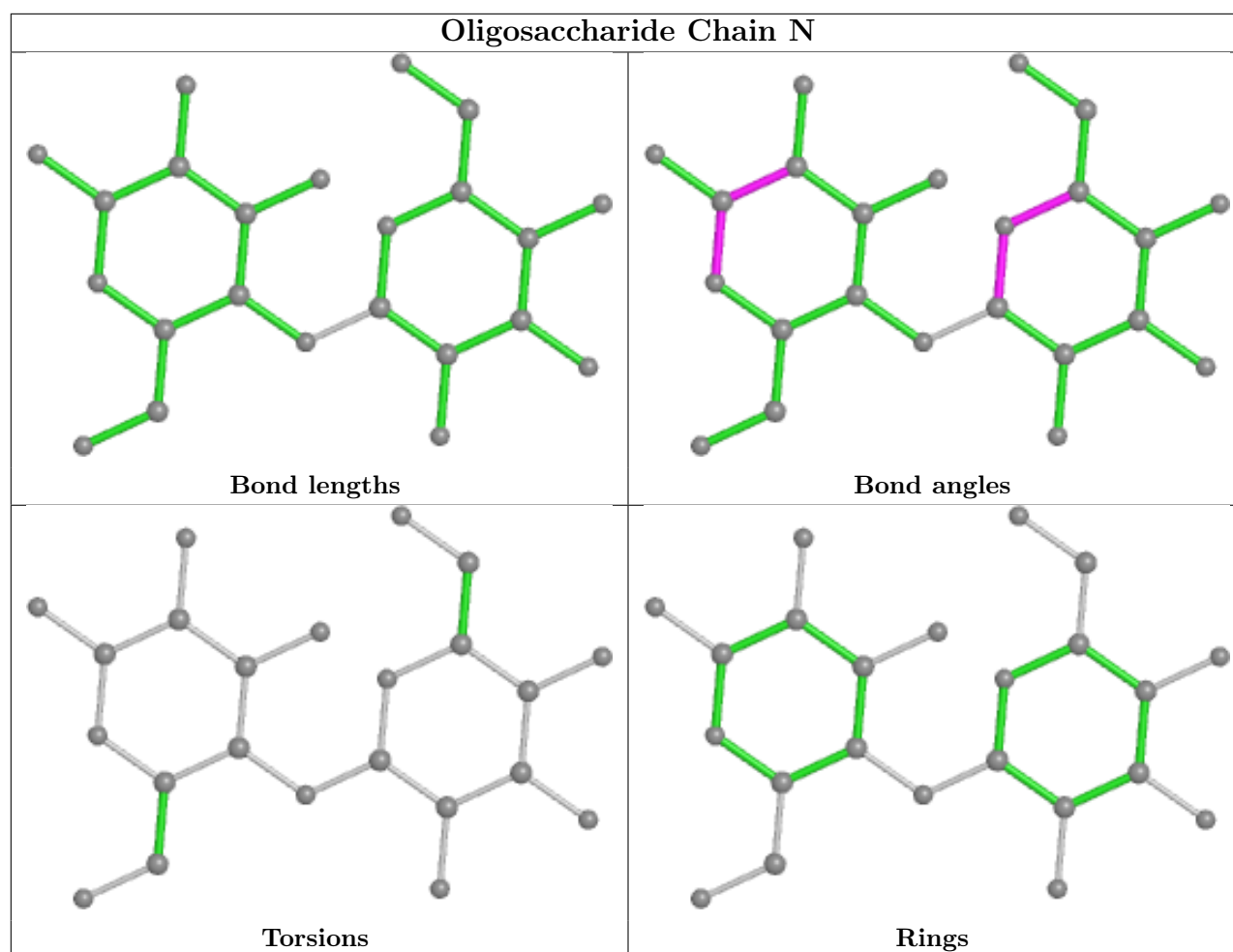


## Oligosaccharide Chain K









## 5.6 Ligand geometry [i](#)

13 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	$\# Z  > 2$
4	EDO	B	102	-	3,3,3	0.42	0	2,2,2	0.32	0
4	EDO	C	104	-	3,3,3	0.65	0	2,2,2	0.30	0
4	EDO	D	103	-	3,3,3	0.53	0	2,2,2	0.38	0
4	EDO	A	103	-	3,3,3	0.44	0	2,2,2	0.44	0
4	EDO	C	101	-	3,3,3	0.48	0	2,2,2	0.38	0
4	EDO	C	102	-	3,3,3	0.47	0	2,2,2	0.40	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EDO	D	104	-	3,3,3	0.57	0	2,2,2	0.18	0
4	EDO	A	101	-	3,3,3	0.49	0	2,2,2	0.30	0
4	EDO	A	102	-	3,3,3	0.46	0	2,2,2	0.35	0
4	EDO	D	102	-	3,3,3	0.50	0	2,2,2	0.33	0
4	EDO	C	103	-	3,3,3	0.50	0	2,2,2	0.33	0
4	EDO	D	101	-	3,3,3	0.50	0	2,2,2	0.41	0
4	EDO	B	101	-	3,3,3	0.51	0	2,2,2	0.22	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
4	EDO	B	102	-	-	1/1/1/1	-
4	EDO	C	104	-	-	1/1/1/1	-
4	EDO	D	103	-	-	0/1/1/1	-
4	EDO	A	103	-	-	1/1/1/1	-
4	EDO	C	101	-	-	0/1/1/1	-
4	EDO	C	102	-	-	0/1/1/1	-
4	EDO	D	104	-	-	1/1/1/1	-
4	EDO	A	101	-	-	0/1/1/1	-
4	EDO	A	102	-	-	1/1/1/1	-
4	EDO	D	102	-	-	0/1/1/1	-
4	EDO	C	103	-	-	0/1/1/1	-
4	EDO	D	101	-	-	1/1/1/1	-
4	EDO	B	101	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	104	EDO	O1-C1-C2-O2
4	D	101	EDO	O1-C1-C2-O2
4	D	104	EDO	O1-C1-C2-O2
4	A	102	EDO	O1-C1-C2-O2
4	B	102	EDO	O1-C1-C2-O2

There are no ring outliers.

7 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	102	EDO	2	0
4	C	104	EDO	3	0
4	A	103	EDO	2	0
4	D	104	EDO	5	0
4	A	102	EDO	2	0
4	D	102	EDO	4	0
4	C	103	EDO	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	436/439 (99%)	0.56	51 (11%) 4 7	25, 41, 98, 106	0
1	B	436/439 (99%)	0.92	79 (18%) 1 1	23, 52, 98, 110	0
1	C	437/439 (99%)	0.04	5 (1%) 80 85	21, 37, 58, 72	0
1	D	437/439 (99%)	0.02	7 (1%) 72 79	23, 38, 57, 72	0
2	E	15/15 (100%)	0.51	1 (6%) 17 26	77, 85, 103, 105	0
2	F	15/15 (100%)	0.66	0 100 100	76, 86, 103, 105	0
2	G	15/15 (100%)	-0.38	0 100 100	30, 33, 42, 46	0
2	H	15/15 (100%)	-0.36	0 100 100	28, 36, 43, 45	0
All	All	1806/1816 (99%)	0.37	143 (7%) 12 19	21, 40, 94, 110	0

The worst 5 of 143 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	27	TRP	9.5
1	B	27	TRP	9.3
1	B	88	ARG	7.9
1	A	57	TYR	6.7
1	B	57	TYR	6.7

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

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

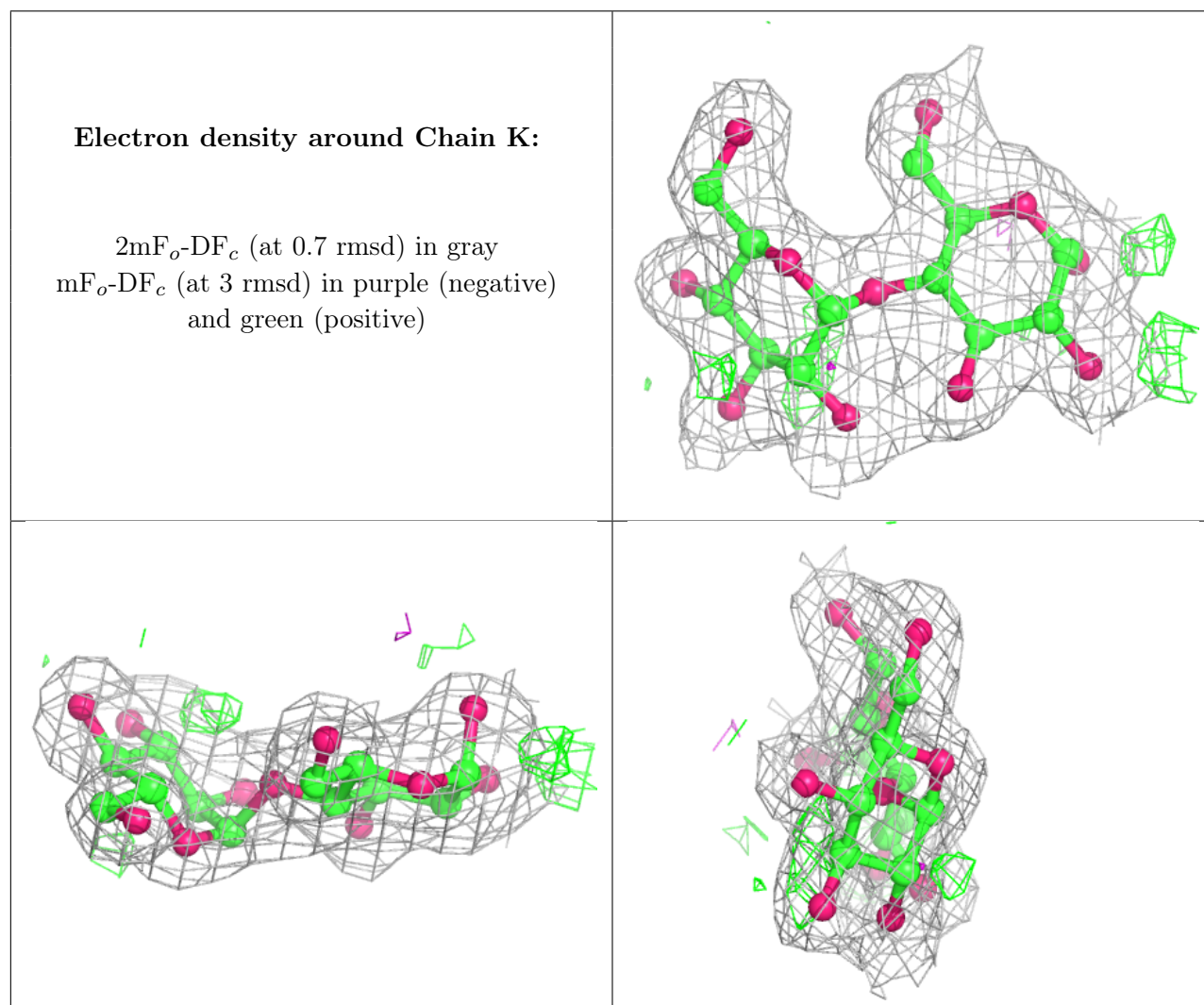
### 6.3 Carbohydrates [i](#)

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
3	GLC	K	2	11/12	0.93	0.09	28,33,37,38	0
3	GLC	K	1	12/12	0.95	0.15	28,34,35,35	0
3	GLC	L	1	12/12	0.96	0.14	32,34,36,39	0
3	GLC	L	2	11/12	0.96	0.13	28,34,37,39	0
3	GLC	M	1	12/12	0.97	0.15	22,26,29,33	0
3	GLC	M	2	11/12	0.98	0.15	20,22,25,26	0
3	GLC	N	1	12/12	0.98	0.14	23,28,30,34	0
3	GLC	N	2	11/12	0.98	0.13	21,23,27,28	0

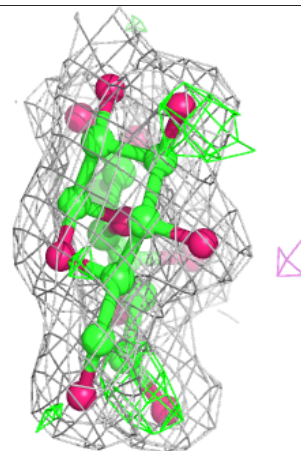
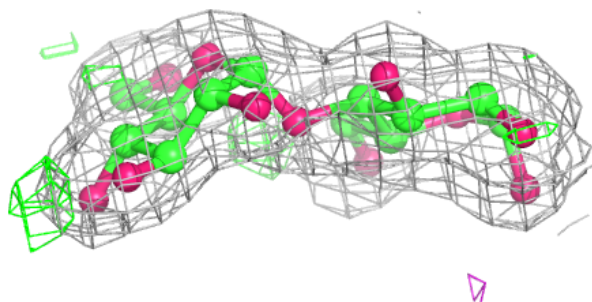
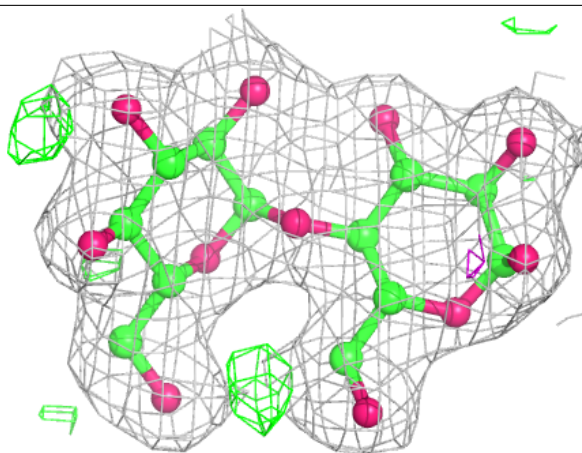
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.





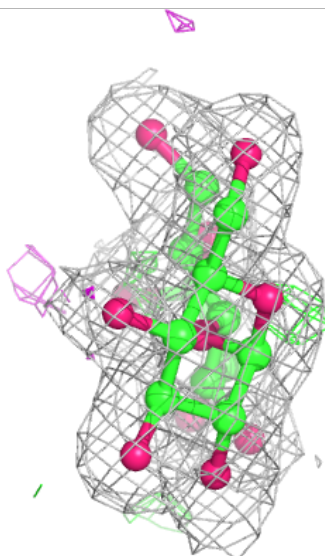
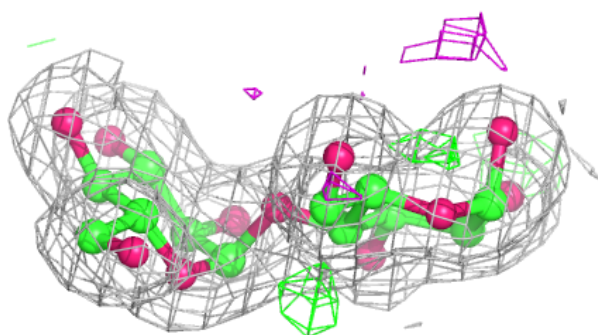
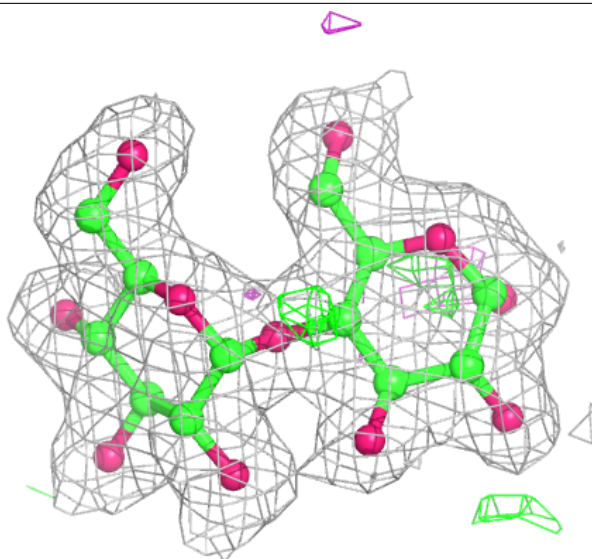
**Electron density around Chain L:**

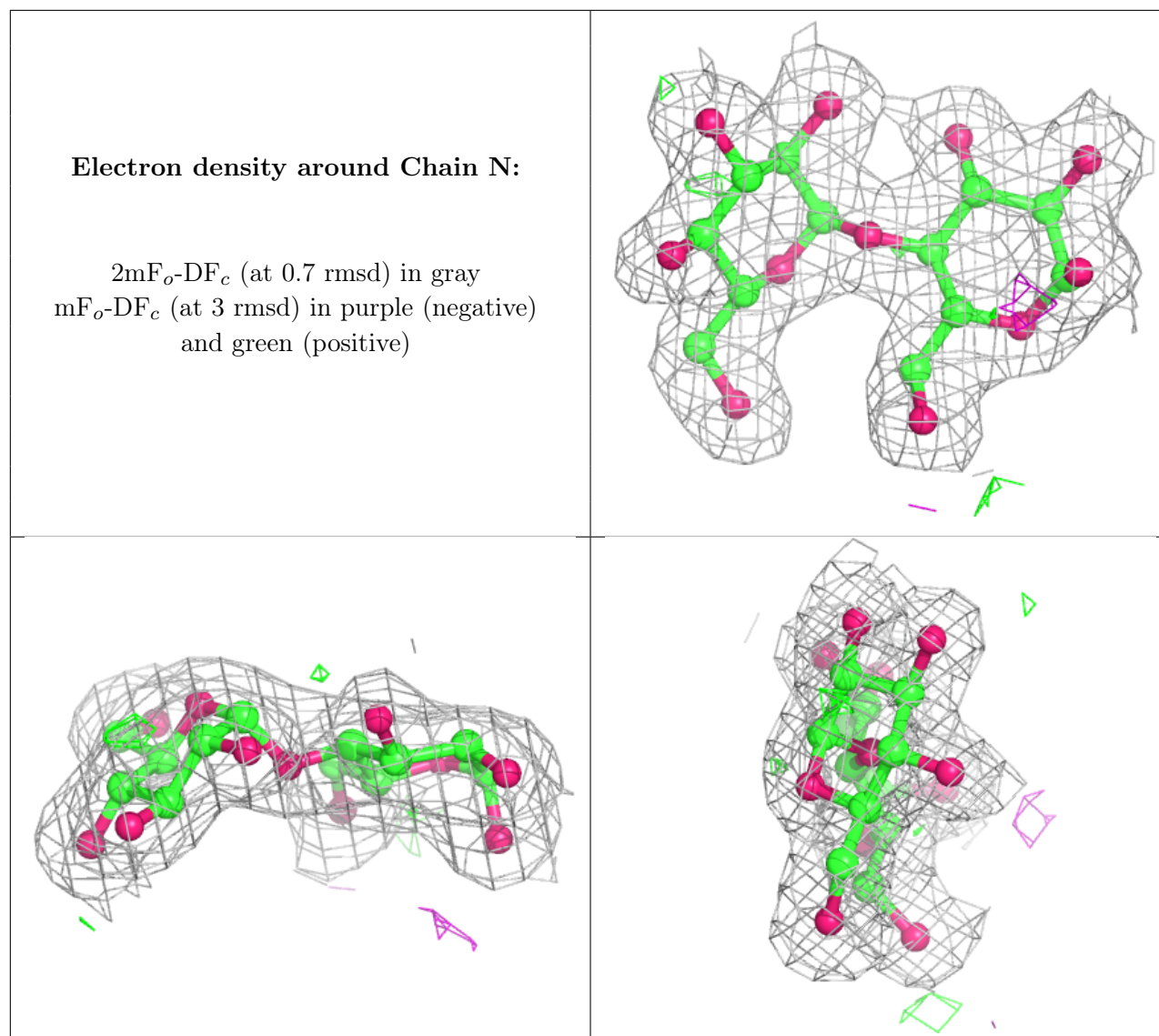
$2mF_o - DF_c$  (at 0.7 rmsd) in gray  
 $mF_o - DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain M:**

$2mF_o-DF_c$  (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
4	EDO	D	101	4/4	0.78	0.21	45,46,46,54	0
4	EDO	C	103	4/4	0.81	0.15	43,45,46,51	0
4	EDO	C	102	4/4	0.83	0.14	44,44,50,54	0
4	EDO	C	104	4/4	0.86	0.21	28,37,50,60	0
4	EDO	D	102	4/4	0.87	0.23	37,45,50,50	0
4	EDO	A	102	4/4	0.88	0.25	39,43,45,46	0
4	EDO	D	104	4/4	0.90	0.20	35,42,55,67	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	EDO	A	103	4/4	0.93	0.36	44,53,61,65	0
4	EDO	D	103	4/4	0.94	0.09	31,38,44,44	0
4	EDO	B	101	4/4	0.94	0.12	36,37,41,42	0
4	EDO	A	101	4/4	0.95	0.12	37,39,39,39	0
4	EDO	B	102	4/4	0.95	0.28	35,42,51,57	0
4	EDO	C	101	4/4	0.96	0.09	31,33,34,38	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.