



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2020 – 10:13 AM BST

PDB ID : 6JWI  
Title : Yeast Npl4 in complex with Lys48-linked diubiquitin  
Authors : Sato, Y.; Fukai, S.  
Deposited on : 2019-04-20  
Resolution : 2.55 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

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

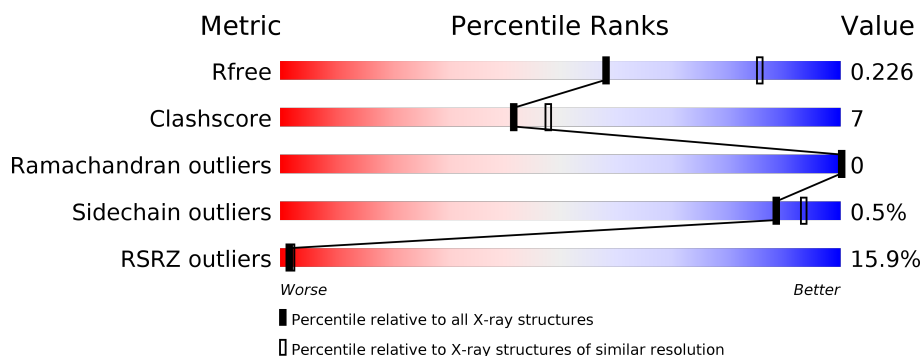
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	G	76	<div> <div>30%</div> <div> <div>72%</div> <div>21%</div> <div>7%</div> </div> </div>
1	I	76	<div> <div>74%</div> <div> <div>72%</div> <div>22%</div> <div>5%</div> </div> </div>
2	A	473	<div> <div>8%</div> <div> <div>82%</div> <div>15%</div> </div> </div>
2	E	473	<div> <div>11%</div> <div> <div>81%</div> <div>15%</div> </div> </div>
3	B	2	<div> <div>50%</div> <div>50%</div> </div>
3	C	2	<div> <div>50%</div> <div>50%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	D	2	 100%
3	F	2	 50%50%
3	H	2	 50%50%
3	J	2	 100%
3	K	2	 100%
3	L	2	 100%
3	M	2	 100%
3	N	2	 50%50%
3	O	2	 100%
3	P	2	 100%
3	Q	2	 100%
3	R	2	 100%

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 9116 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 Ubiquitin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	72	Total	C	N	O	Se	0	0	0
			574	362	98	113	1			
1	G	71	Total	C	N	O	Se	0	0	0
			563	356	94	112	1			

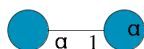
- Molecule 2 is a protein called Nuclear protein localization protein 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	459	Total	C	N	O	S	0	0	0
			3692	2339	616	716	21			
2	E	459	Total	C	N	O	S	0	0	0
			3692	2339	616	716	21			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	108	GLY	-	expression tag	UNP P33755
A	109	PRO	-	expression tag	UNP P33755
A	110	LEU	-	expression tag	UNP P33755
A	111	GLY	-	expression tag	UNP P33755
A	112	SER	-	expression tag	UNP P33755
E	108	GLY	-	expression tag	UNP P33755
E	109	PRO	-	expression tag	UNP P33755
E	110	LEU	-	expression tag	UNP P33755
E	111	GLY	-	expression tag	UNP P33755
E	112	SER	-	expression tag	UNP P33755

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

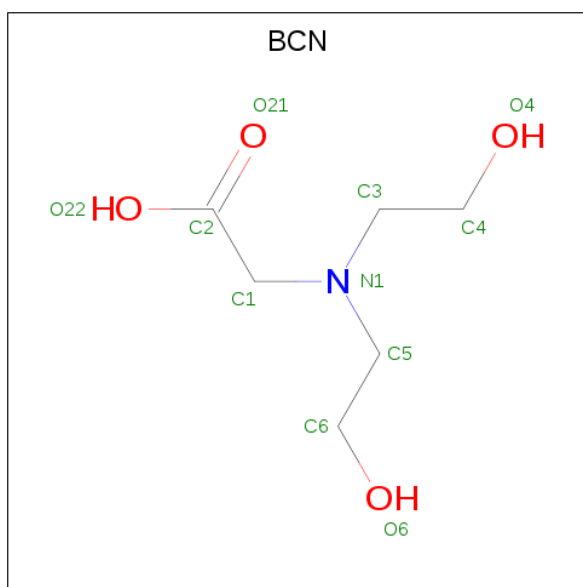


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
3	B	2	Total	C	O	0	0	0
			23	12	11			
3	C	2	Total	C	O	0	0	0
			23	12	11			
3	D	2	Total	C	O	0	0	0
			23	12	11			
3	F	2	Total	C	O	0	0	0
			23	12	11			
3	H	2	Total	C	O	0	0	0
			23	12	11			
3	J	2	Total	C	O	0	0	0
			23	12	11			
3	K	2	Total	C	O	0	0	0
			23	12	11			
3	L	2	Total	C	O	0	0	0
			23	12	11			
3	M	2	Total	C	O	0	0	0
			23	12	11			
3	N	2	Total	C	O	0	0	0
			23	12	11			
3	O	2	Total	C	O	0	0	0
			23	12	11			
3	P	2	Total	C	O	0	0	0
			23	12	11			
3	Q	2	Total	C	O	0	0	0
			23	12	11			
3	R	2	Total	C	O	0	0	0
			23	12	11			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Zn	0	0
			2	2		
4	E	2	Total	Zn	0	0
			2	2		

- Molecule 5 is BICINE (three-letter code: BCN) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			11	6	1	4		
5	A	1	Total	C	N	O	0	0
			11	6	1	4		
5	A	1	Total	C	N	O	0	0
			11	6	1	4		
5	A	1	Total	C	N	O	0	0
			11	6	1	4		
5	E	1	Total	C	N	O	0	0
			11	6	1	4		
5	E	1	Total	C	N	O	0	0
			11	6	1	4		
5	E	1	Total	C	N	O	0	0
			11	6	1	4		

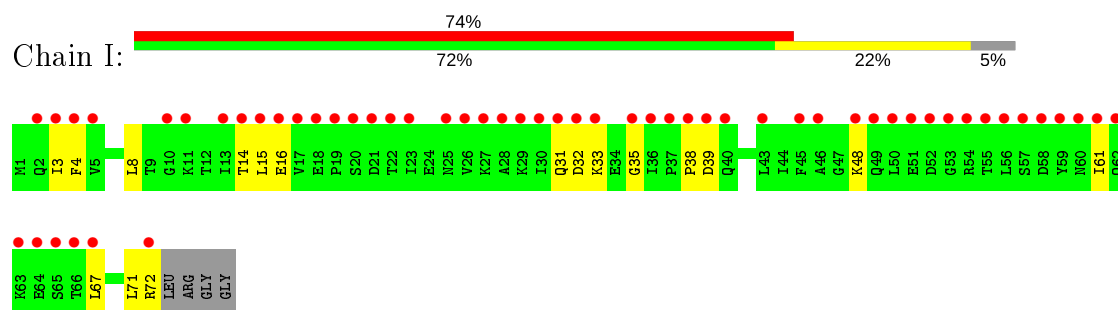
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	2	Total	O	0	0
			2	2		
6	A	79	Total	O	0	0
			79	79		
6	E	89	Total	O	0	0
			89	89		

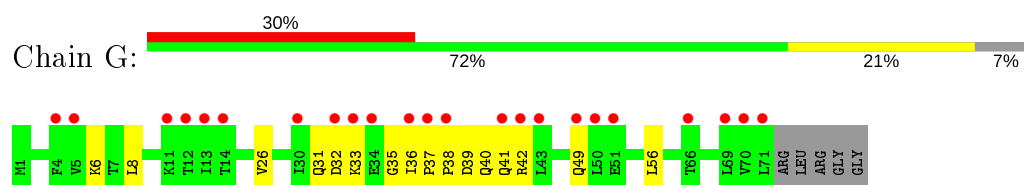
### 3 Residue-property plots [i](#)

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.

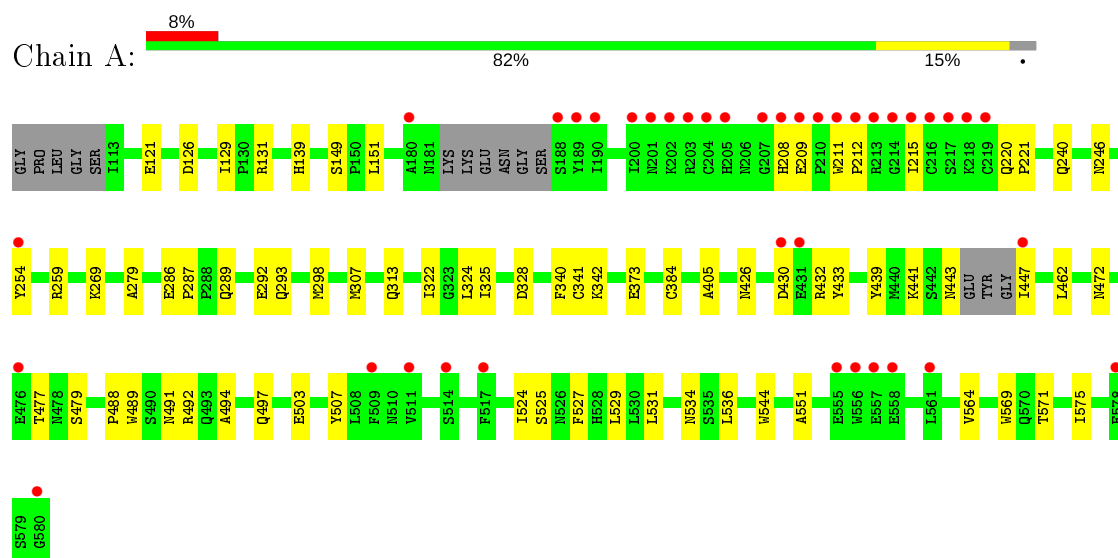
#### • Molecule 1: Ubiquitin



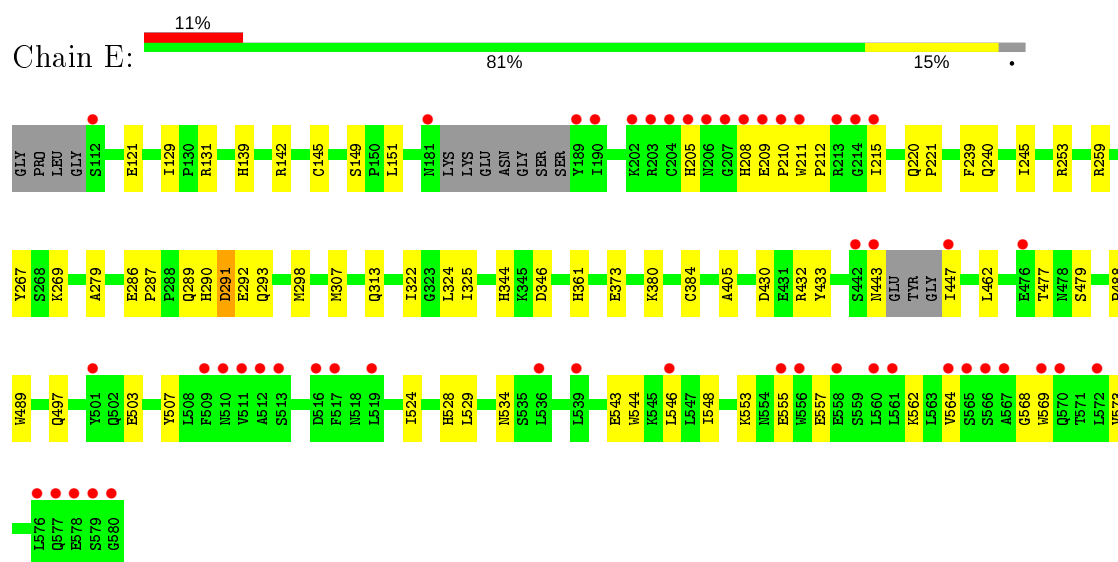
#### • Molecule 1: Ubiquitin



#### • Molecule 2: Nuclear protein localization protein 4



#### • Molecule 2: Nuclear protein localization protein 4



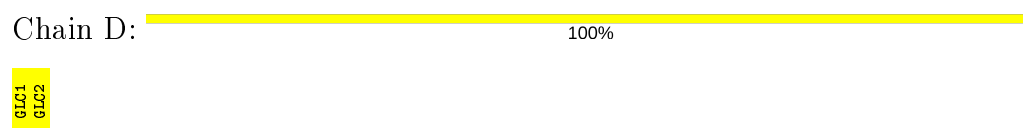
- Molecule 3: alpha-D-glucopyranose-(1-1)-alpha-D-glucopyranose



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



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



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



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



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



Chain J:  100%

GLC1  
GLC2

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

Chain K:  100%

GLC1  
GLC2

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

Chain L:  100%

GLC1  
GLC2

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

Chain M:  100%

GLC1  
GLC2

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

Chain N:  50% 50%

GLC1  
GLC2

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

Chain O:  100%

GLC1  
GLC2

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

Chain P:  100%

GLC1  
GLC2

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

Chain Q:  100%

GLC1  
1271  
GLC2

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

Chain R:

100%

GLC1  
1272  
GLC2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.35Å 103.10Å 99.62Å 90.00° 100.40° 90.00°	Depositor
Resolution (Å)	48.99 – 2.55 48.99 – 2.55	Depositor EDS
% Data completeness (in resolution range)	98.9 (48.99-2.55) 98.9 (48.99-2.55)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.25	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.58 (at 2.54Å)	Xtriage
Refinement program	PHENIX 1.16_3549	Depositor
R, $R_{free}$	0.191 , 0.226 0.190 , 0.226	Depositor DCC
$R_{free}$ test set	2831 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.3	Xtriage
Anisotropy	0.426	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 58.1	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	0.94	EDS
Total number of atoms	9116	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.89% 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: BCN, GLC, ZN

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	G	0.31	0/569	0.55	0/767
1	I	0.29	0/580	0.56	0/781
2	A	0.47	0/3778	0.63	0/5102
2	E	0.46	0/3778	0.63	0/5102
All	All	0.45	0/8705	0.62	0/11752

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	563	0	586	12	0
1	I	574	0	599	12	0
2	A	3692	0	3566	48	0
2	E	3692	0	3566	46	1
3	B	23	0	21	0	0
3	C	23	0	21	1	0
3	D	23	0	21	0	0
3	F	23	0	21	1	0
3	H	23	0	21	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	J	23	0	21	2	0
3	K	23	0	21	0	0
3	L	23	0	21	2	0
3	M	23	0	21	1	0
3	N	23	0	21	0	0
3	O	23	0	21	2	0
3	P	23	0	21	0	0
3	Q	23	0	21	0	0
3	R	23	0	21	1	0
4	A	2	0	0	0	0
4	E	2	0	0	0	0
5	A	55	0	60	4	0
5	E	44	0	48	7	0
6	A	79	0	0	1	0
6	E	89	0	0	2	0
6	I	2	0	0	0	0
All	All	9116	0	8719	124	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

The worst 5 of 124 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)
2:A:536:LEU:O	1:G:6:LYS:NZ	2.12	0.82
2:E:212:PRO:HG3	2:E:430:ASP:HB2	1.63	0.81
2:A:477:THR:HG22	2:A:479:SER:H	1.47	0.80
2:A:341:CYS:HB3	5:A:613:BCN:H52	1.68	0.76
1:I:39:ASP:O	1:I:72:ARG:NH1	2.22	0.73

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:142:ARG:NH2	2:E:253:ARG:O[2_646]	2.17	0.03

## 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	G	69/76 (91%)	68 (99%)	1 (1%)	0	100	100
1	I	70/76 (92%)	69 (99%)	1 (1%)	0	100	100
2	A	453/473 (96%)	444 (98%)	9 (2%)	0	100	100
2	E	453/473 (96%)	442 (98%)	11 (2%)	0	100	100
All	All	1045/1098 (95%)	1023 (98%)	22 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	65/67 (97%)	65 (100%)	0	100	100
1	I	66/67 (98%)	65 (98%)	1 (2%)	65	77
2	A	413/423 (98%)	412 (100%)	1 (0%)	93	97
2	E	413/423 (98%)	410 (99%)	3 (1%)	84	90
All	All	957/980 (98%)	952 (100%)	5 (0%)	88	93

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	16	GLU
2	A	373	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	E	291	ASP
2	E	373	GLU
2	E	432	ARG

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

Mol	Chain	Res	Type
1	I	40	GLN
2	E	240	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

28 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	B	1	3	11,11,12	0.72	0	15,15,17	0.95	0
3	GLC	B	2	3	12,12,12	1.06	2 (16%)	17,17,17	0.93	0
3	GLC	C	1	3	11,11,12	0.62	0	15,15,17	1.10	1 (6%)
3	GLC	C	2	3	12,12,12	1.09	1 (8%)	17,17,17	0.95	2 (11%)
3	GLC	D	1	3	11,11,12	0.85	0	15,15,17	1.55	3 (20%)
3	GLC	D	2	3	12,12,12	1.26	1 (8%)	17,17,17	1.09	1 (5%)
3	GLC	F	1	3	11,11,12	0.95	0	15,15,17	1.75	5 (33%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GLC	F	2	3	12,12,12	1.00	0	17,17,17	0.68	0
3	GLC	H	1	3	11,11,12	1.24	1 (9%)	15,15,17	0.72	0
3	GLC	H	2	3	12,12,12	1.23	2 (16%)	17,17,17	0.85	0
3	GLC	J	1	3	11,11,12	1.33	2 (18%)	15,15,17	1.43	2 (13%)
3	GLC	J	2	3	12,12,12	1.44	2 (16%)	17,17,17	1.53	5 (29%)
3	GLC	K	1	3	11,11,12	0.97	0	15,15,17	1.14	1 (6%)
3	GLC	K	2	3	12,12,12	1.32	2 (16%)	17,17,17	1.49	3 (17%)
3	GLC	L	1	3	11,11,12	1.38	2 (18%)	15,15,17	1.20	2 (13%)
3	GLC	L	2	3	12,12,12	1.43	2 (16%)	17,17,17	1.04	2 (11%)
3	GLC	M	1	3	11,11,12	1.01	1 (9%)	15,15,17	1.28	3 (20%)
3	GLC	M	2	3	12,12,12	0.66	0	17,17,17	0.87	0
3	GLC	N	1	3	11,11,12	0.79	0	15,15,17	0.88	1 (6%)
3	GLC	N	2	3	12,12,12	0.71	0	17,17,17	0.76	0
3	GLC	O	1	3	11,11,12	1.26	1 (9%)	15,15,17	2.07	5 (33%)
3	GLC	O	2	3	12,12,12	1.85	4 (33%)	17,17,17	1.97	6 (35%)
3	GLC	P	1	3	11,11,12	2.08	6 (54%)	15,15,17	1.35	2 (13%)
3	GLC	P	2	3	12,12,12	1.33	2 (16%)	17,17,17	1.51	5 (29%)
3	GLC	Q	1	3	11,11,12	0.95	0	15,15,17	0.79	0
3	GLC	Q	2	3	12,12,12	0.84	0	17,17,17	0.95	0
3	GLC	R	1	3	11,11,12	1.54	2 (18%)	15,15,17	2.11	6 (40%)
3	GLC	R	2	3	12,12,12	1.36	1 (8%)	17,17,17	1.32	3 (17%)

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	B	1	3	-	2/2/19/22	0/1/1/1
3	GLC	B	2	3	-	2/2/22/22	0/1/1/1
3	GLC	C	1	3	-	0/2/19/22	0/1/1/1
3	GLC	C	2	3	-	2/2/22/22	0/1/1/1
3	GLC	D	1	3	-	0/2/19/22	0/1/1/1
3	GLC	D	2	3	-	2/2/22/22	0/1/1/1
3	GLC	F	1	3	-	2/2/19/22	0/1/1/1
3	GLC	F	2	3	-	2/2/22/22	0/1/1/1
3	GLC	H	1	3	-	0/2/19/22	0/1/1/1

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GLC	H	2	3	-	1/2/22/22	0/1/1/1
3	GLC	J	1	3	-	2/2/19/22	0/1/1/1
3	GLC	J	2	3	-	2/2/22/22	0/1/1/1
3	GLC	K	1	3	-	2/2/19/22	0/1/1/1
3	GLC	K	2	3	-	2/2/22/22	0/1/1/1
3	GLC	L	1	3	-	1/2/19/22	0/1/1/1
3	GLC	L	2	3	-	2/2/22/22	0/1/1/1
3	GLC	M	1	3	-	1/2/19/22	0/1/1/1
3	GLC	M	2	3	-	0/2/22/22	0/1/1/1
3	GLC	N	1	3	-	0/2/19/22	0/1/1/1
3	GLC	N	2	3	-	2/2/22/22	0/1/1/1
3	GLC	O	1	3	-	0/2/19/22	0/1/1/1
3	GLC	O	2	3	-	2/2/22/22	0/1/1/1
3	GLC	P	1	3	-	1/2/19/22	0/1/1/1
3	GLC	P	2	3	-	2/2/22/22	0/1/1/1
3	GLC	Q	1	3	-	2/2/19/22	0/1/1/1
3	GLC	Q	2	3	-	0/2/22/22	0/1/1/1
3	GLC	R	1	3	-	0/2/19/22	0/1/1/1
3	GLC	R	2	3	-	2/2/22/22	0/1/1/1

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	O	2	GLC	C1-C2	4.49	1.63	1.52
3	P	1	GLC	C1-C2	3.79	1.60	1.52
3	R	1	GLC	O5-C5	3.70	1.50	1.43
3	R	2	GLC	O1-C1	3.52	1.50	1.39
3	O	1	GLC	C1-C2	3.48	1.60	1.52

The worst 5 of 58 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	O	2	GLC	C4-C3-C2	-5.19	101.77	110.82
3	R	1	GLC	O5-C5-C6	5.15	115.28	107.20
3	O	1	GLC	O2-C2-C1	4.71	118.79	109.15
3	K	2	GLC	O5-C5-C6	4.00	116.37	106.44
3	P	2	GLC	O5-C5-C6	3.85	116.00	106.44

There are no chirality outliers.

5 of 36 torsion outliers are listed below:

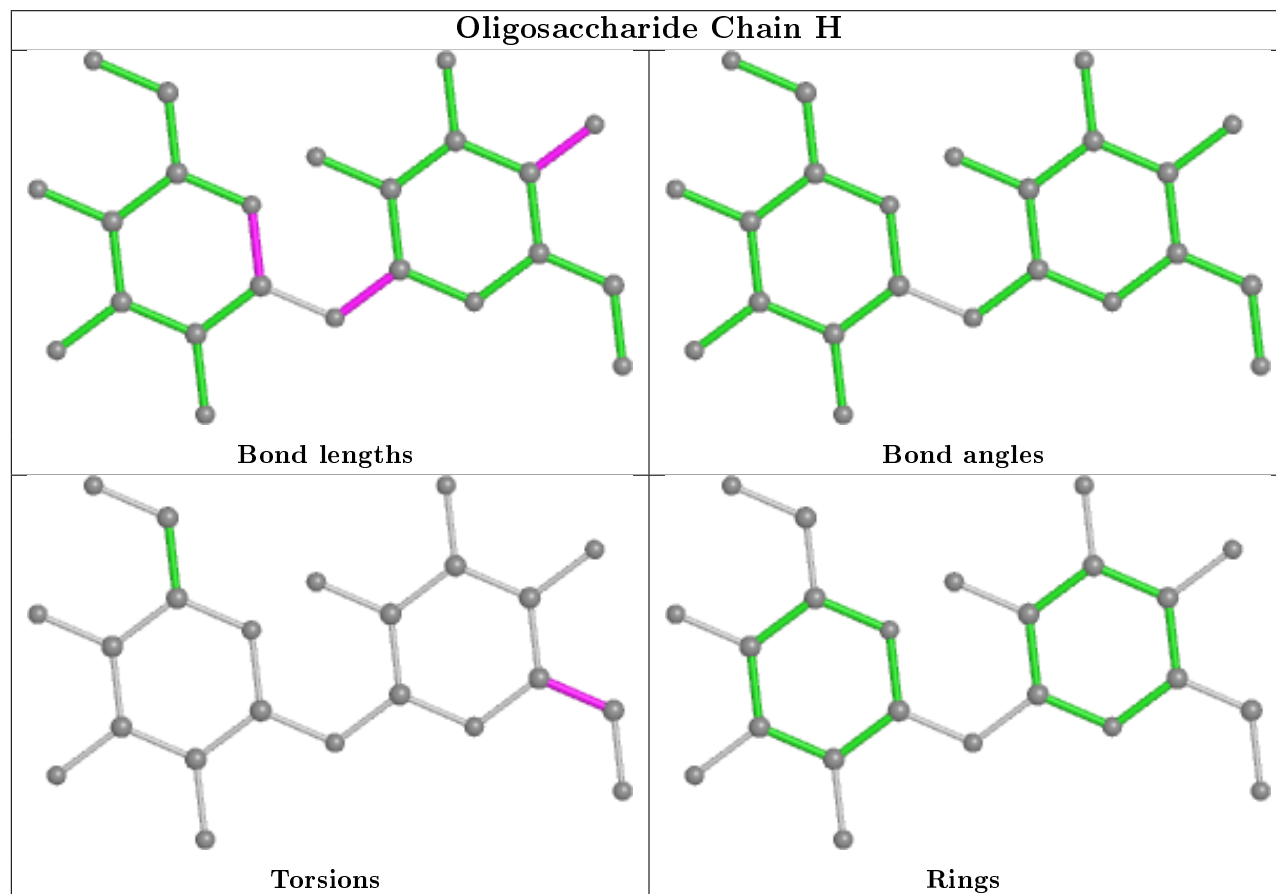
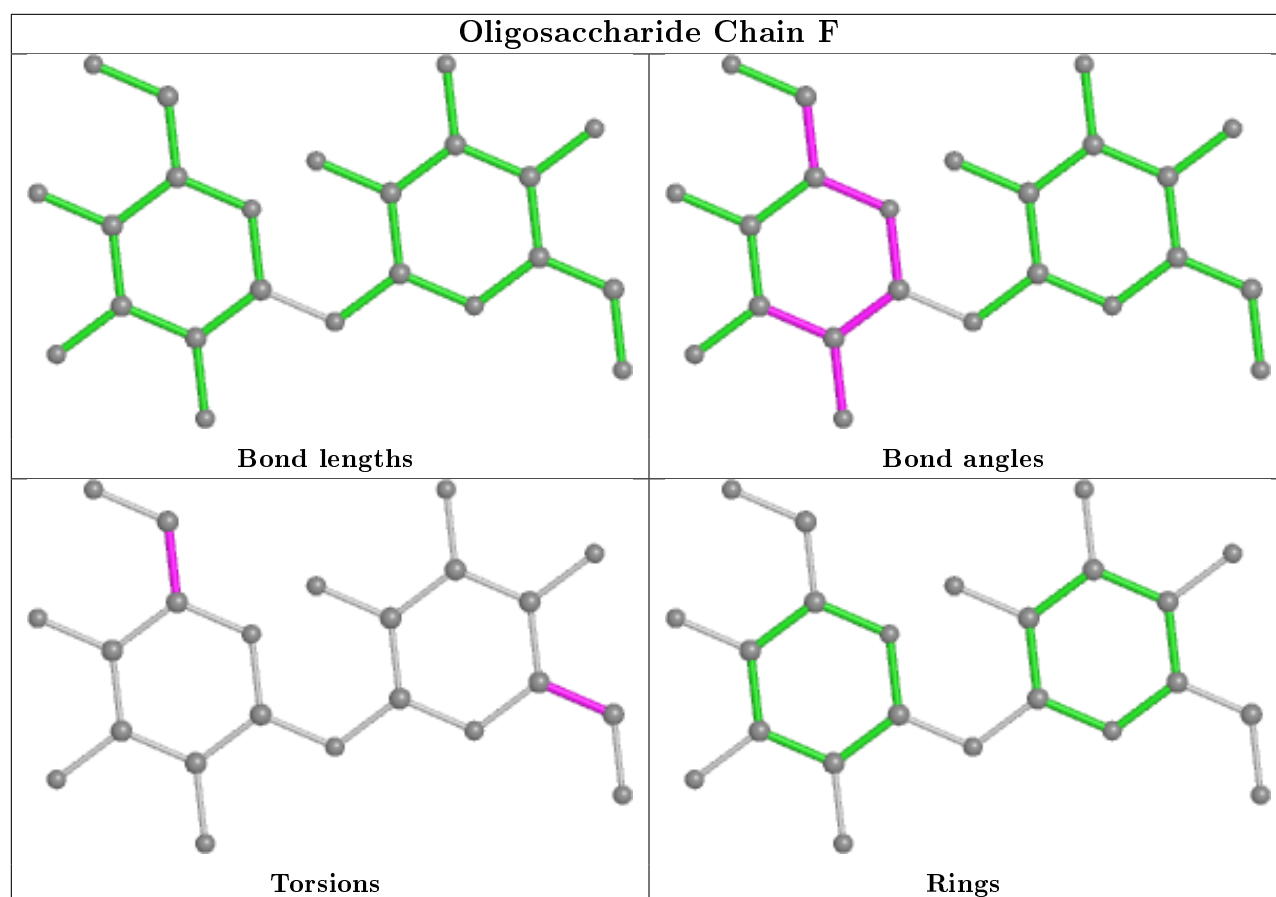
Mol	Chain	Res	Type	Atoms
3	B	2	GLC	C4-C5-C6-O6
3	N	2	GLC	C4-C5-C6-O6
3	R	2	GLC	O5-C5-C6-O6
3	F	2	GLC	O5-C5-C6-O6
3	K	2	GLC	O5-C5-C6-O6

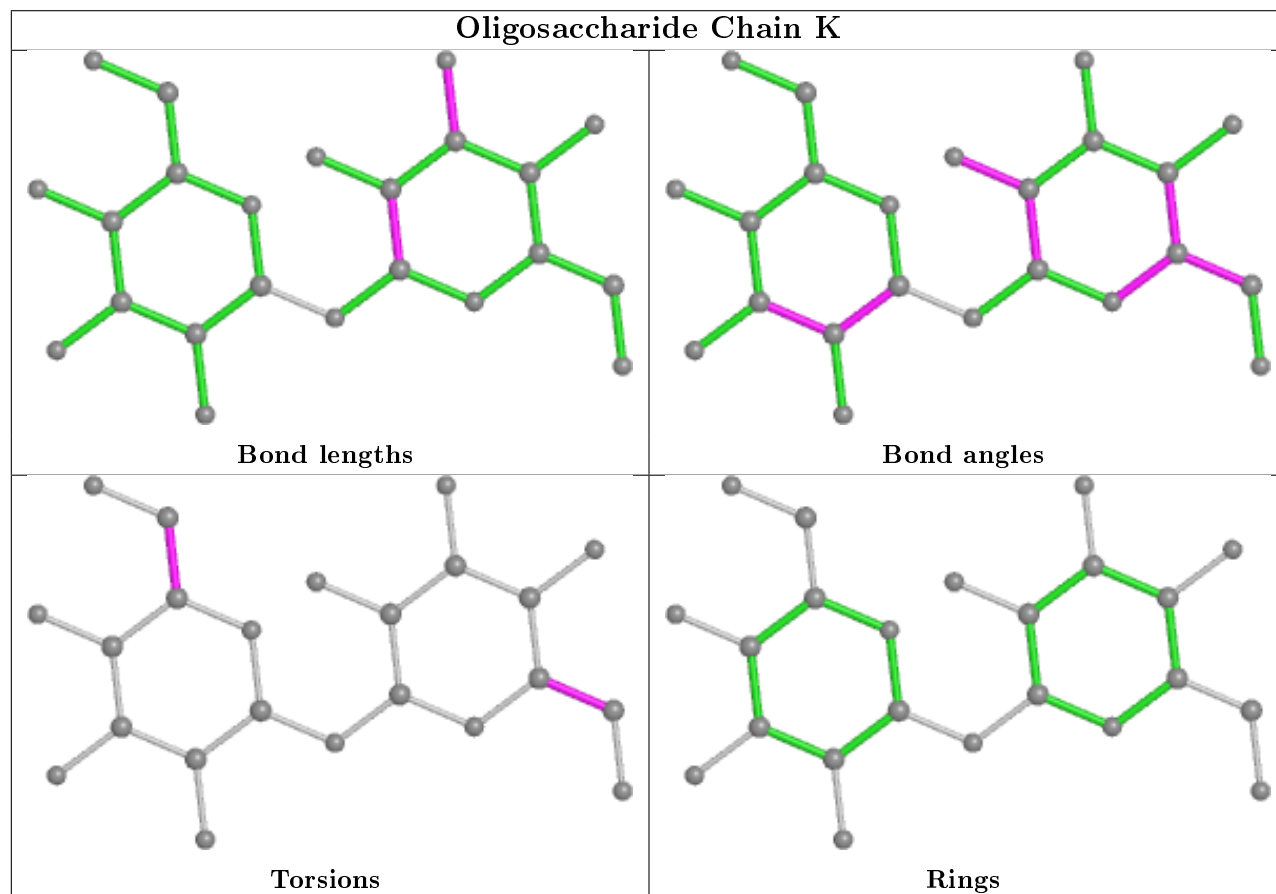
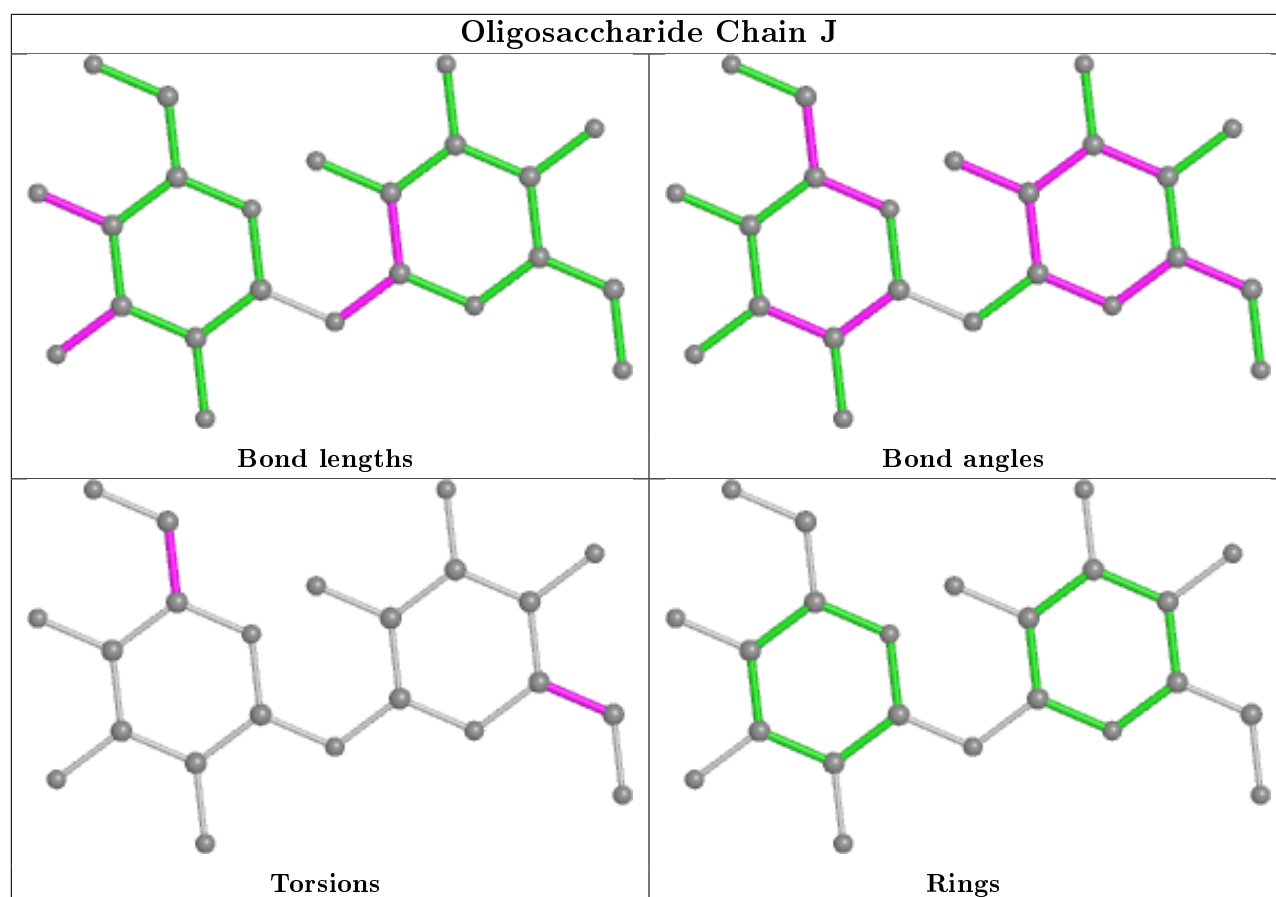
There are no ring outliers.

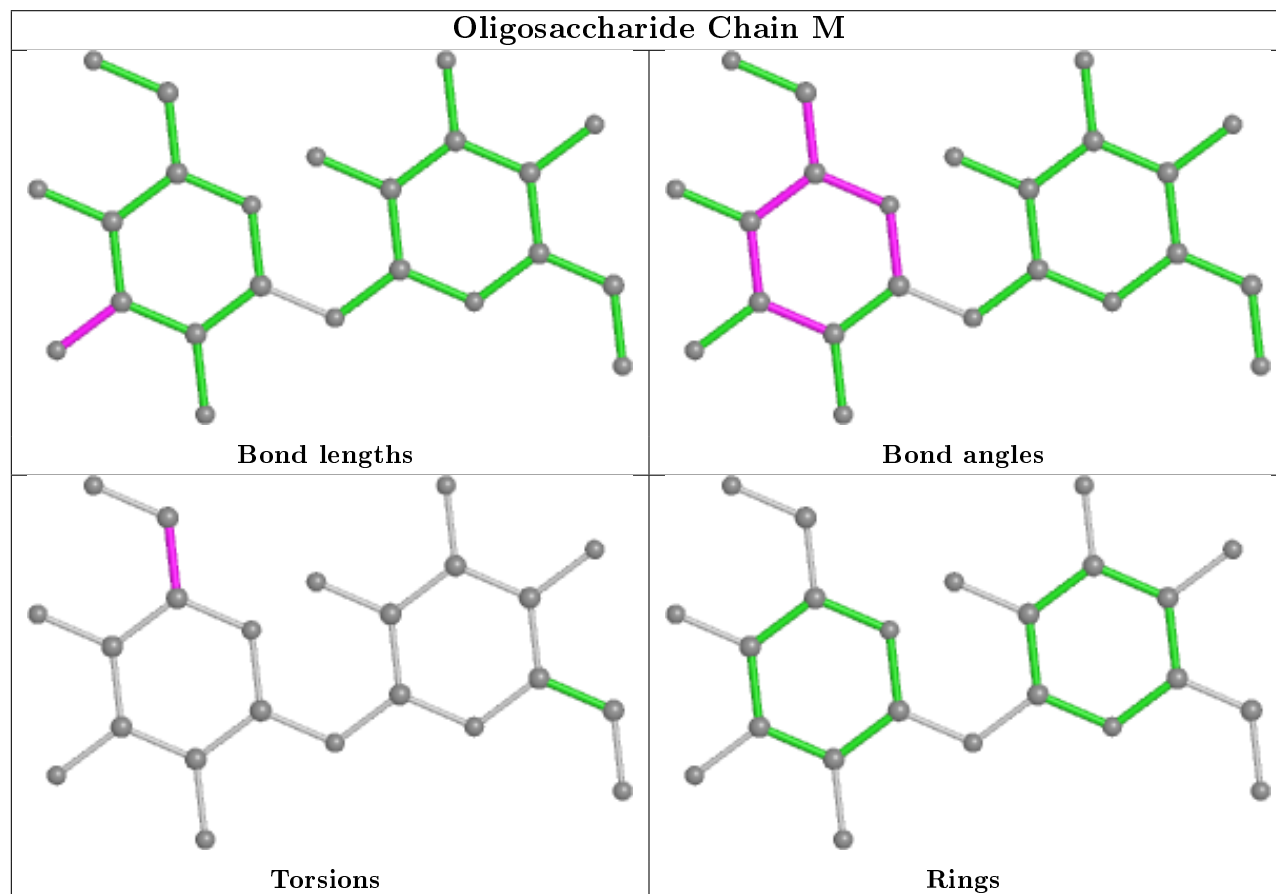
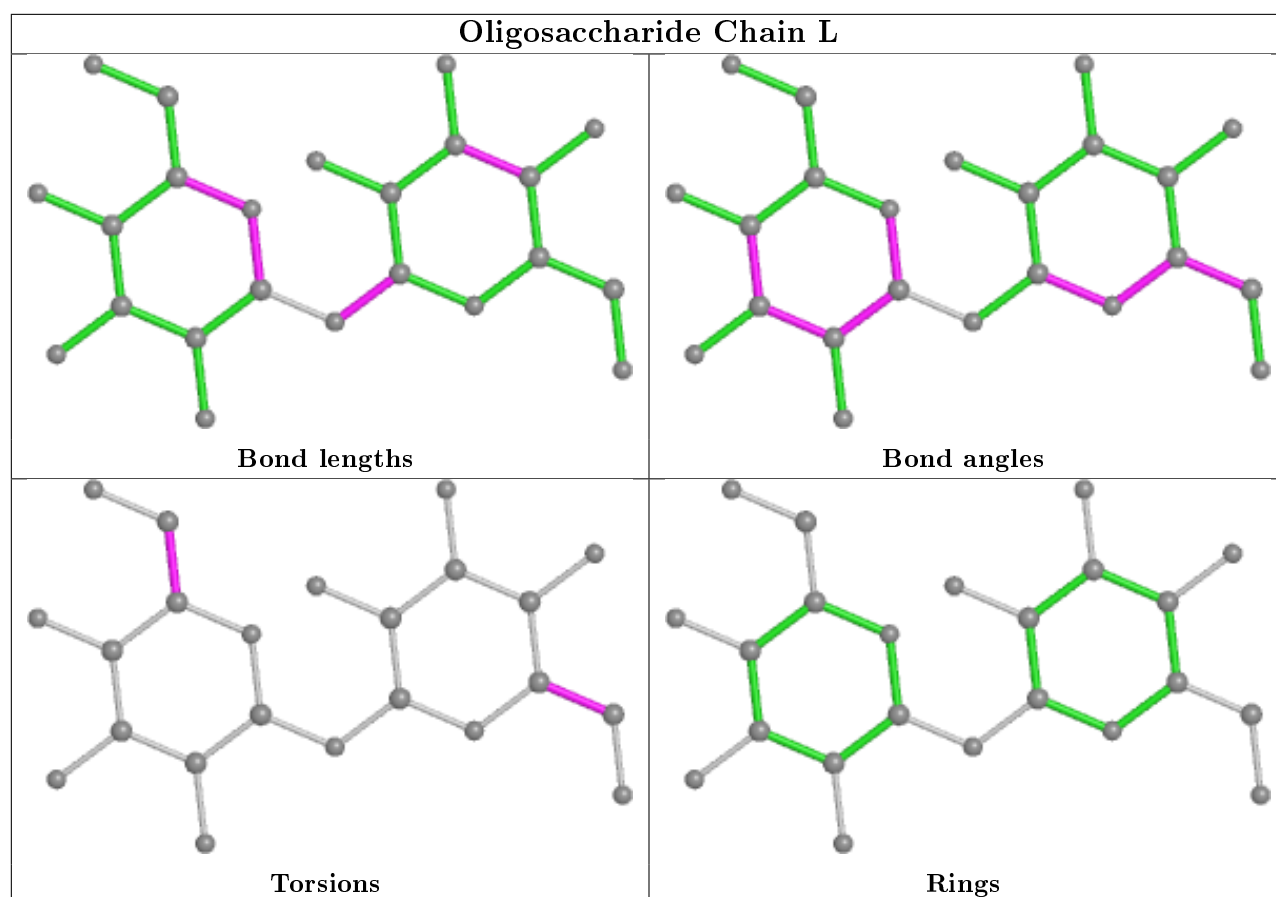
12 monomers are involved in 12 short contacts:

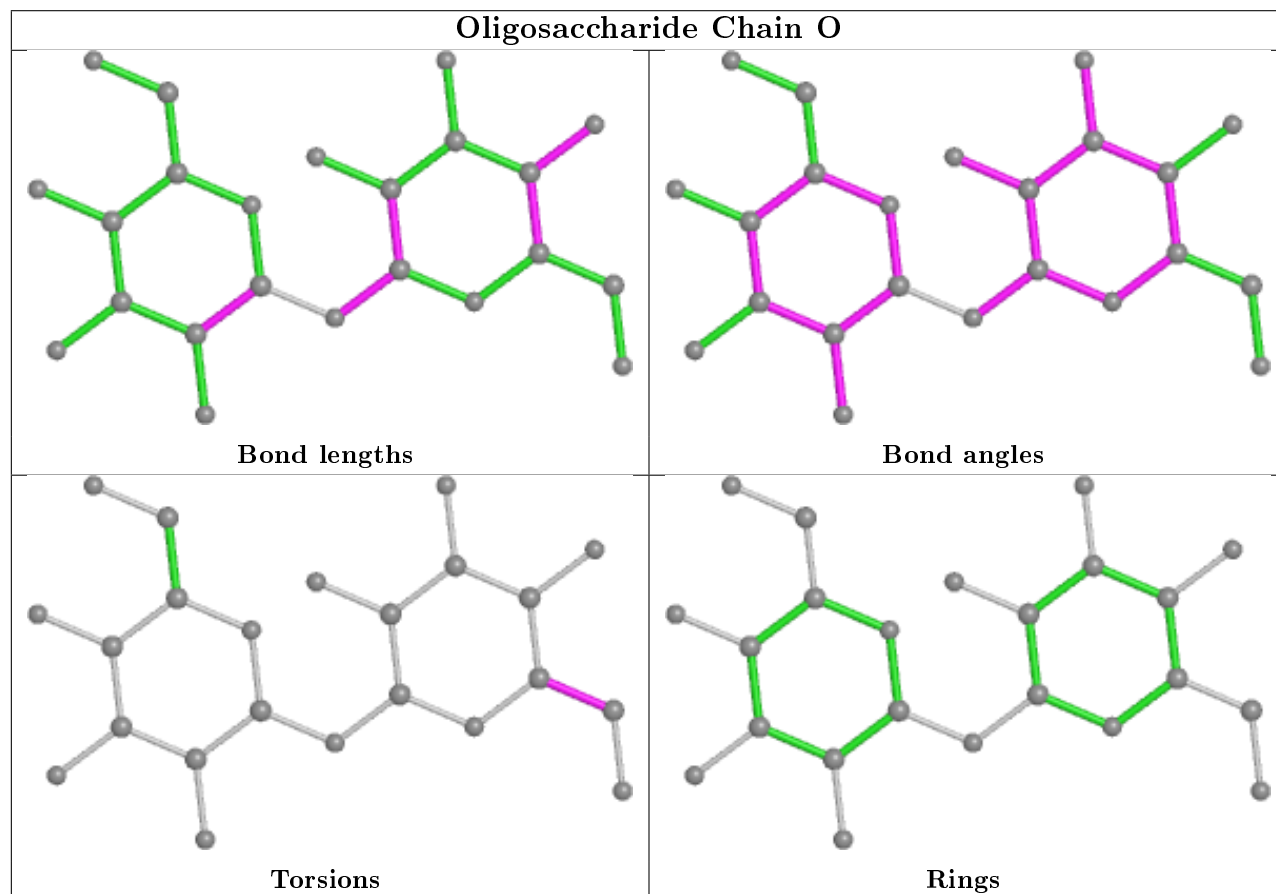
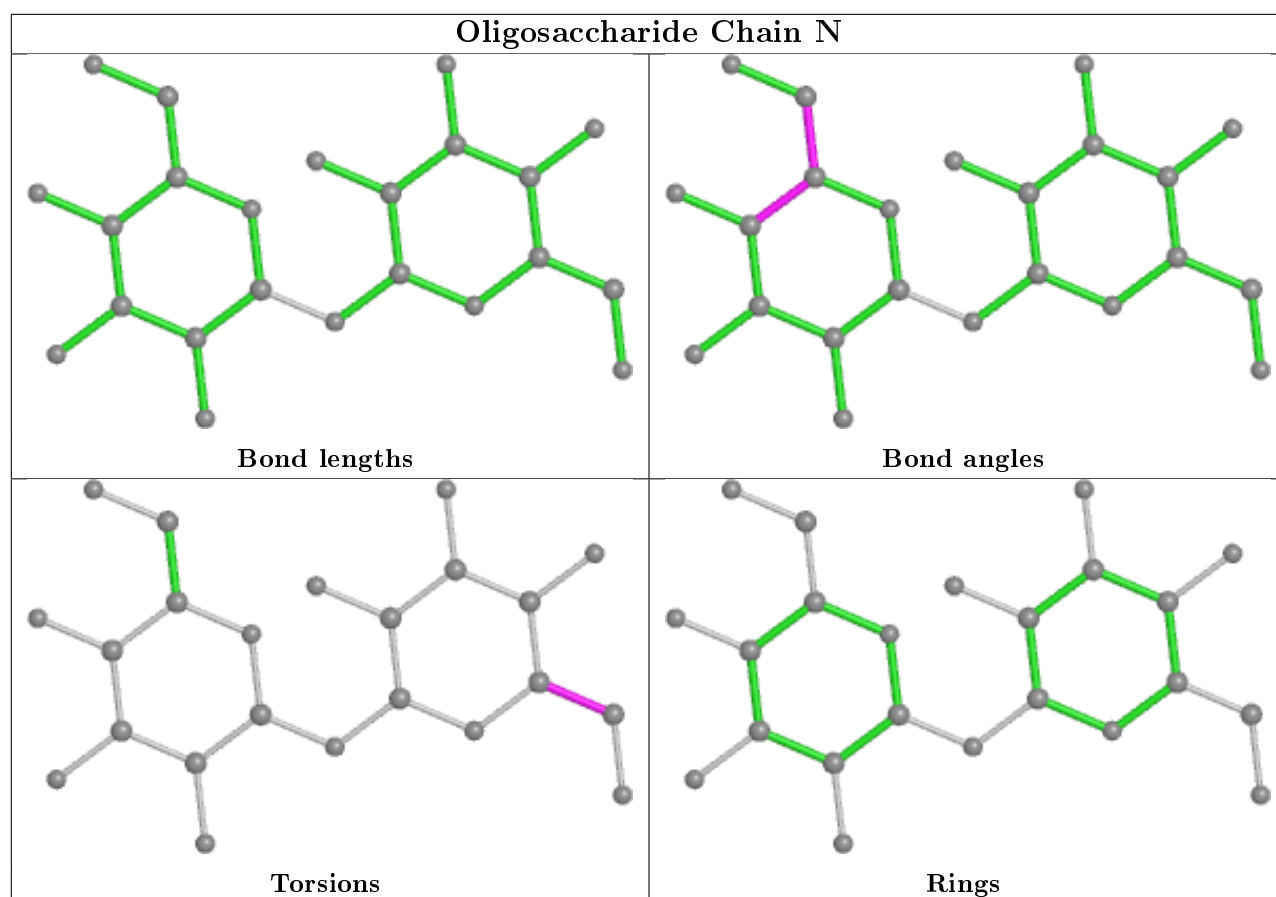
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	M	2	GLC	1	0
3	J	1	GLC	1	0
3	R	1	GLC	1	0
3	F	1	GLC	1	0
3	H	2	GLC	3	0
3	R	2	GLC	1	0
3	L	1	GLC	2	0
3	C	1	GLC	1	0
3	L	2	GLC	1	0
3	O	1	GLC	1	0
3	J	2	GLC	1	0
3	O	2	GLC	1	0

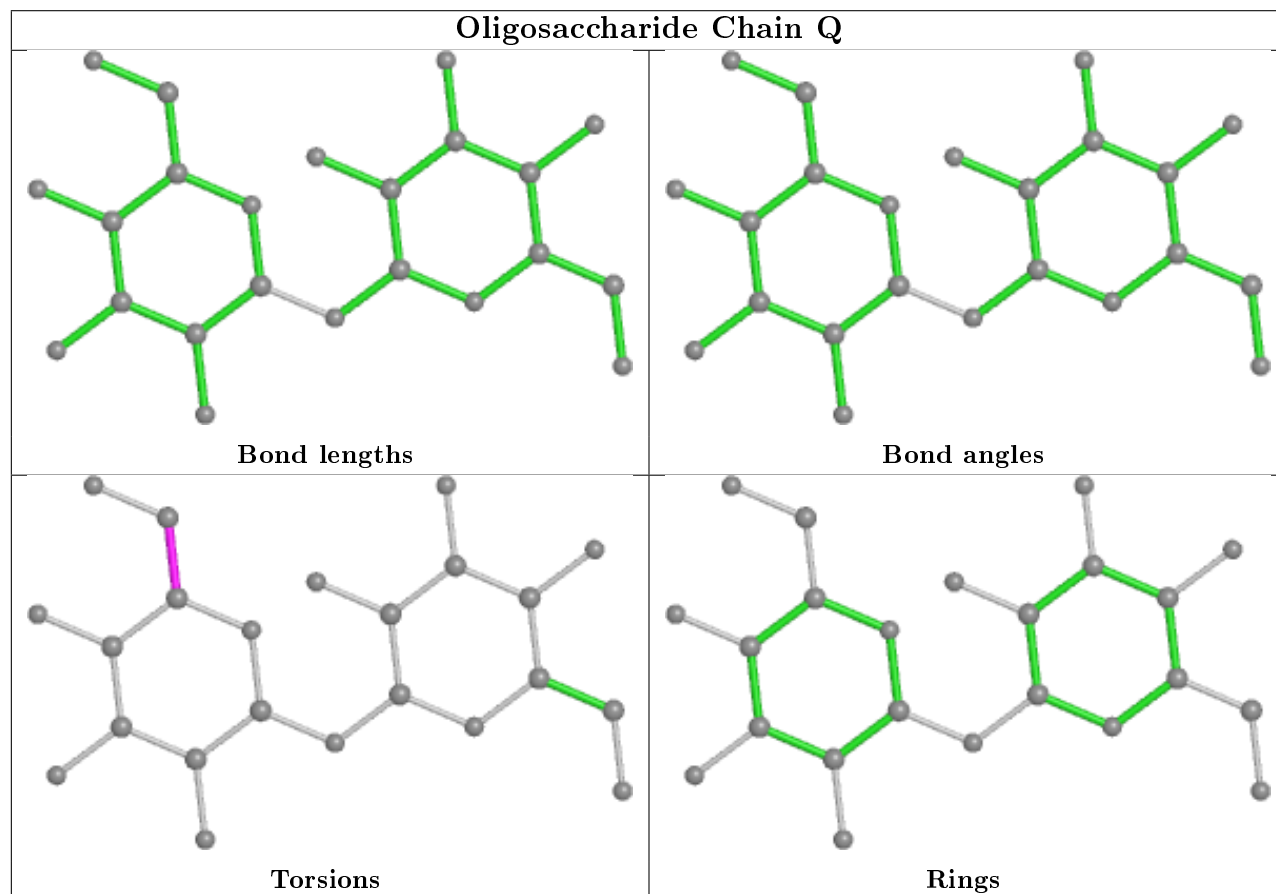
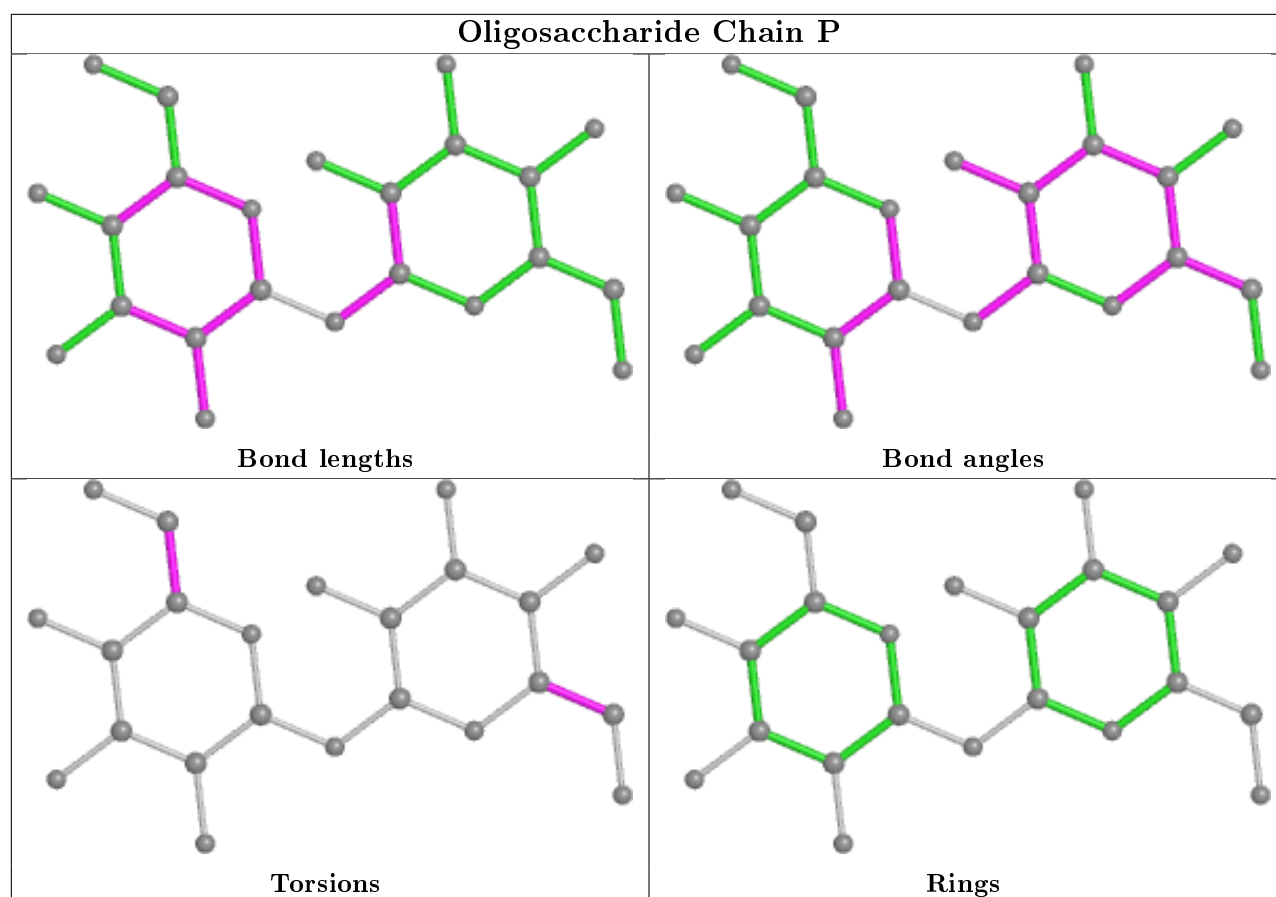
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

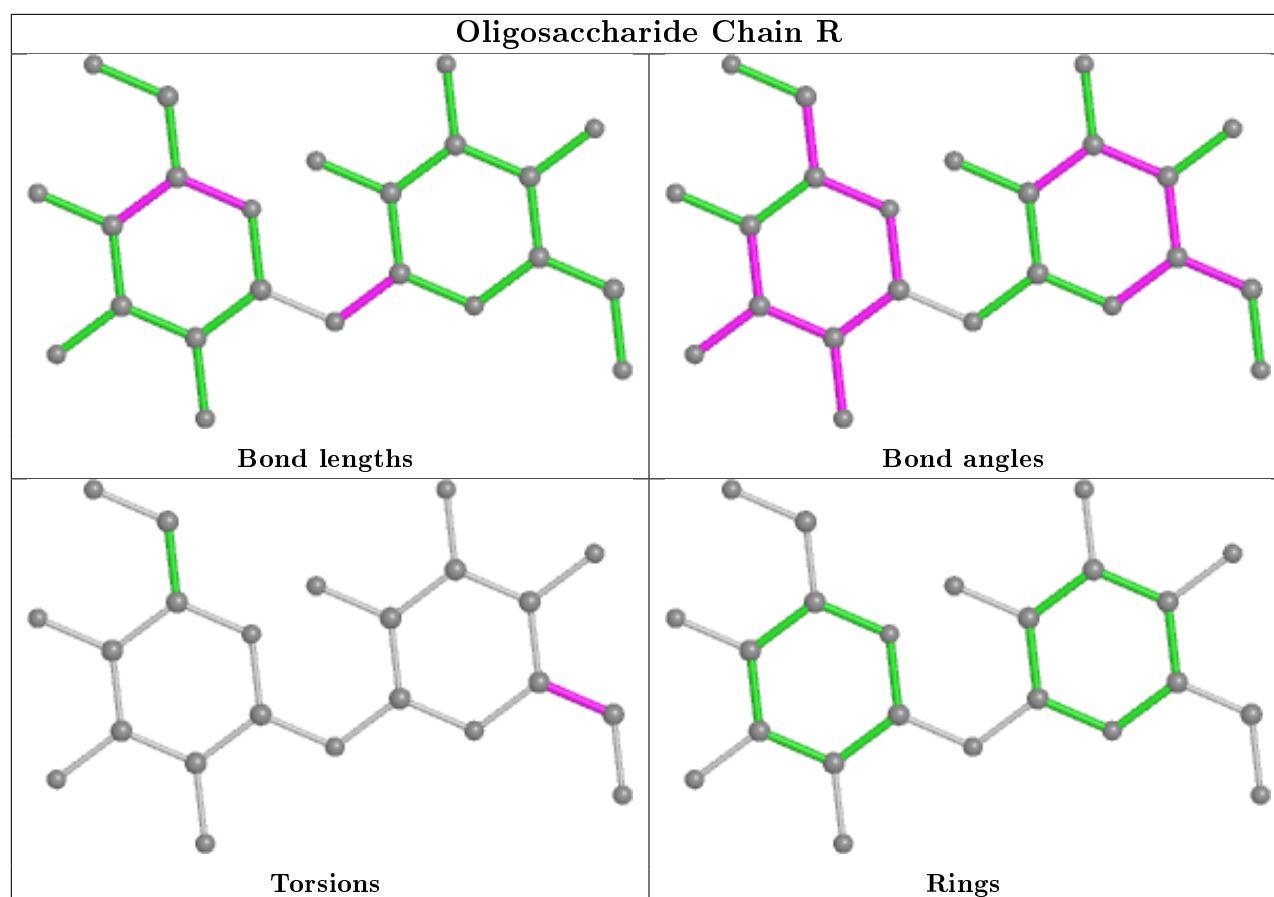












## 5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 4 are monoatomic - leaving 9 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	BCN	E	612	-	7,10,10	0.35	0	8,11,11	0.70	0
5	BCN	E	611	-	7,10,10	0.66	0	8,11,11	1.18	0
5	BCN	A	613	-	7,10,10	0.43	0	8,11,11	0.48	0
5	BCN	E	610	-	7,10,10	0.35	0	8,11,11	0.89	0
5	BCN	E	609	-	7,10,10	0.44	0	8,11,11	0.71	0
5	BCN	A	611	-	7,10,10	0.33	0	8,11,11	0.86	0
5	BCN	A	614	-	7,10,10	0.45	0	8,11,11	0.56	0
5	BCN	A	615	-	7,10,10	0.45	0	8,11,11	1.17	1 (12%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	BCN	A	612	-	7,10,10	0.73	0	8,11,11	1.83	2 (25%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	BCN	E	612	-	-	3/8/10/10	-
5	BCN	E	611	-	-	3/8/10/10	-
5	BCN	A	613	-	-	2/8/10/10	-
5	BCN	E	610	-	-	4/8/10/10	-
5	BCN	E	609	-	-	4/8/10/10	-
5	BCN	A	611	-	-	5/8/10/10	-
5	BCN	A	614	-	-	4/8/10/10	-
5	BCN	A	615	-	-	2/8/10/10	-
5	BCN	A	612	-	-	5/8/10/10	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	612	BCN	C2-C1-N1	4.56	119.97	113.48
5	A	615	BCN	C2-C1-N1	-2.95	109.28	113.48
5	A	612	BCN	C1-N1-C5	2.02	115.56	111.29

There are no chirality outliers.

5 of 32 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	610	BCN	N1-C5-C6-O6
5	E	609	BCN	N1-C5-C6-O6
5	A	612	BCN	C4-C3-N1-C5
5	A	612	BCN	C6-C5-N1-C1
5	A	612	BCN	N1-C5-C6-O6

There are no ring outliers.

6 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	612	BCN	4	0
5	E	611	BCN	1	0
5	A	613	BCN	3	0
5	E	610	BCN	1	0
5	E	609	BCN	1	0
5	A	612	BCN	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	G	70/76 (92%)	1.71	23 (32%) 0 0	70, 101, 133, 150	0
1	I	71/76 (93%)	4.34	56 (78%) 0 0	82, 133, 162, 191	0
2	A	459/473 (97%)	0.41	39 (8%) 10 12	19, 41, 94, 154	0
2	E	459/473 (97%)	0.58	50 (10%) 5 7	20, 40, 100, 157	0
All	All	1059/1098 (96%)	0.83	168 (15%) 1 2	19, 46, 127, 191	0

The worst 5 of 168 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	26	VAL	12.2
2	E	580	GLY	10.8
1	I	17	VAL	8.9
1	I	2	GLN	8.8
1	I	15	LEU	8.7

### 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	J	1	11/12	0.65	0.39	81,96,105,107	0
3	GLC	P	1	11/12	0.65	0.27	83,96,101,103	0

*Continued on next page...*

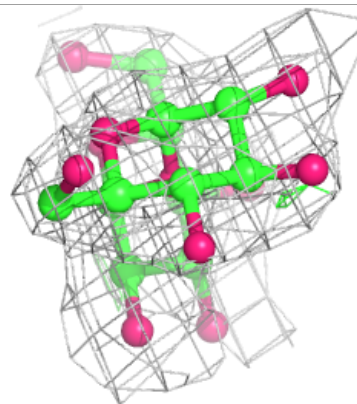
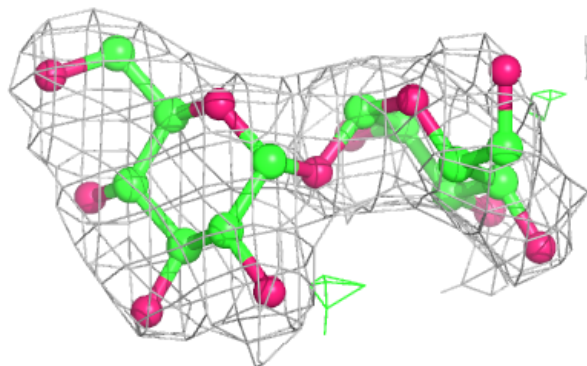
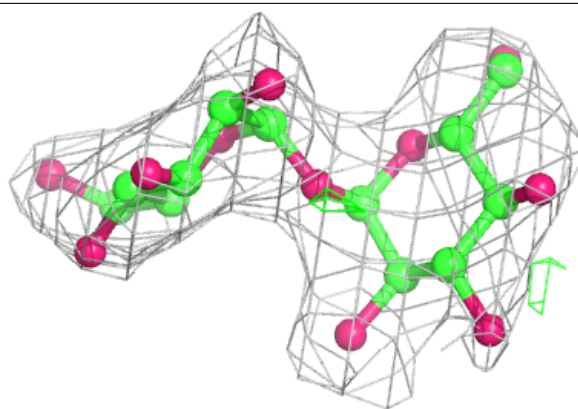
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	GLC	L	2	12/12	0.71	0.28	69,98,104,105	0
3	GLC	D	2	12/12	0.74	0.32	85,94,95,97	0
3	GLC	R	2	12/12	0.75	0.29	78,90,96,97	0
3	GLC	J	2	12/12	0.77	0.26	99,106,114,115	0
3	GLC	R	1	11/12	0.82	0.26	52,65,88,90	0
3	GLC	K	2	12/12	0.82	0.26	61,69,82,85	0
3	GLC	P	2	12/12	0.82	0.23	80,85,88,89	0
3	GLC	D	1	11/12	0.83	0.23	70,76,85,87	0
3	GLC	O	2	12/12	0.84	0.24	50,69,81,84	0
3	GLC	H	1	11/12	0.88	0.15	74,82,86,88	0
3	GLC	F	2	12/12	0.89	0.18	57,72,81,83	0
3	GLC	L	1	11/12	0.90	0.21	68,82,91,93	0
3	GLC	H	2	12/12	0.91	0.34	63,73,78,85	0
3	GLC	Q	2	12/12	0.92	0.17	49,65,68,68	0
3	GLC	Q	1	11/12	0.92	0.26	72,76,85,87	0
3	GLC	K	1	11/12	0.92	0.23	39,67,75,79	0
3	GLC	O	1	11/12	0.93	0.15	49,56,61,62	0
3	GLC	C	2	12/12	0.93	0.17	40,44,49,55	0
3	GLC	B	1	11/12	0.93	0.13	45,52,55,60	0
3	GLC	B	2	12/12	0.94	0.10	50,62,68,71	0
3	GLC	N	1	11/12	0.94	0.10	43,47,59,65	0
3	GLC	F	1	11/12	0.95	0.16	53,56,60,61	0
3	GLC	N	2	12/12	0.95	0.13	59,66,75,76	0
3	GLC	M	1	11/12	0.96	0.21	34,41,44,48	0
3	GLC	C	1	11/12	0.96	0.13	32,39,44,47	0
3	GLC	M	2	12/12	0.97	0.16	31,37,44,44	0

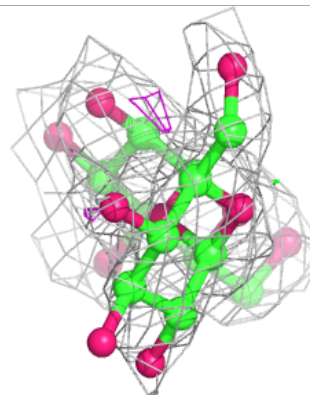
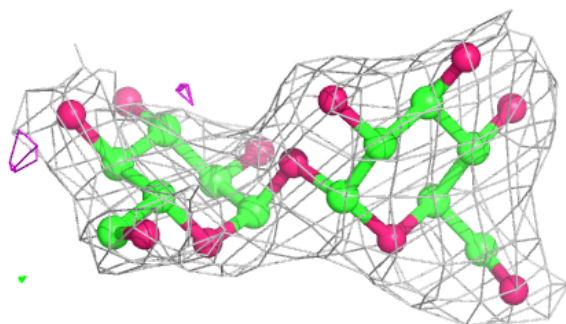
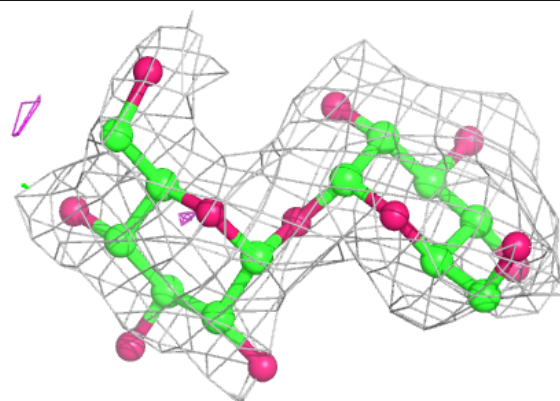
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain F:**

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

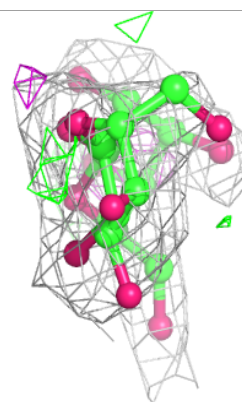
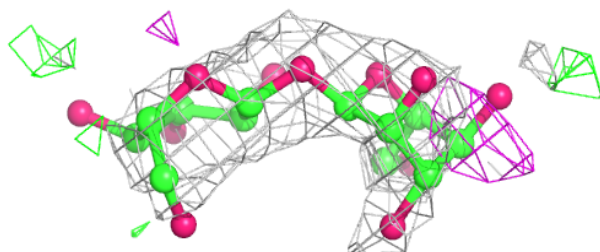
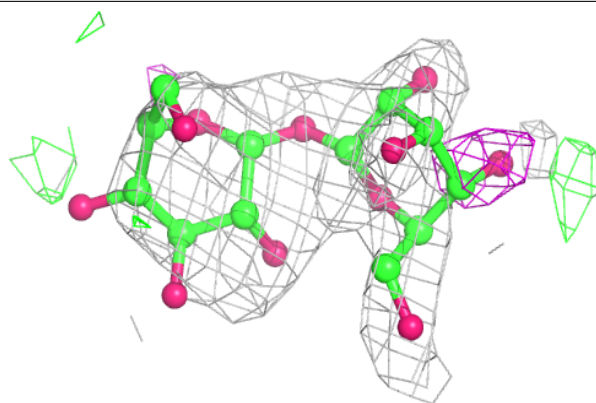
**Electron density around Chain H:**

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

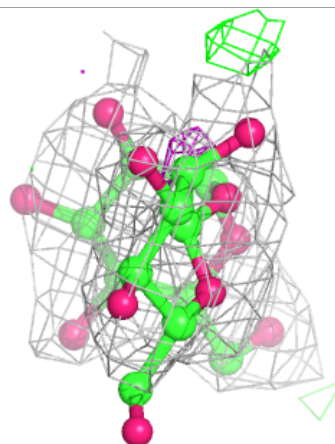
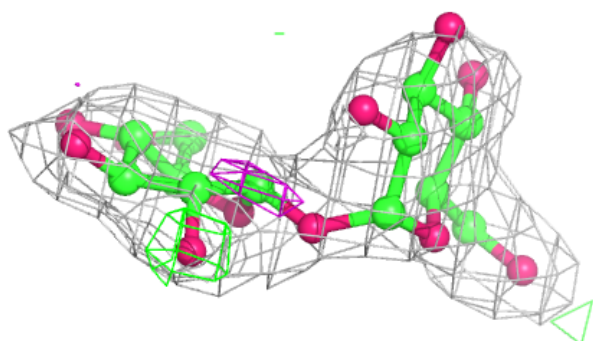
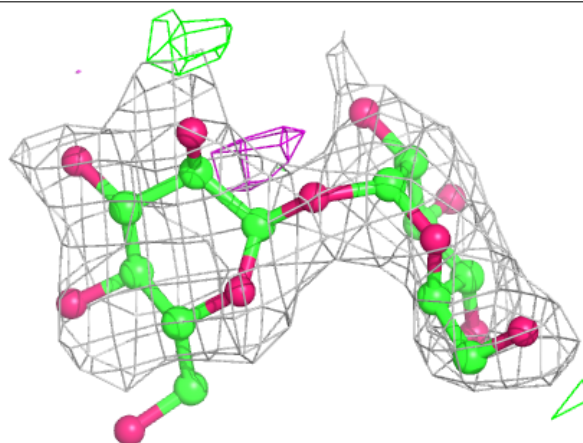


**Electron density around Chain J:**

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

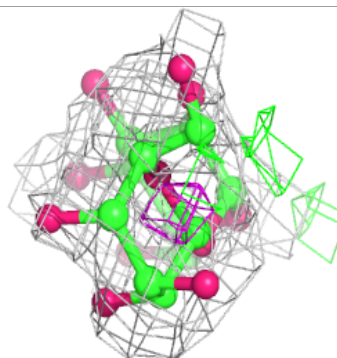
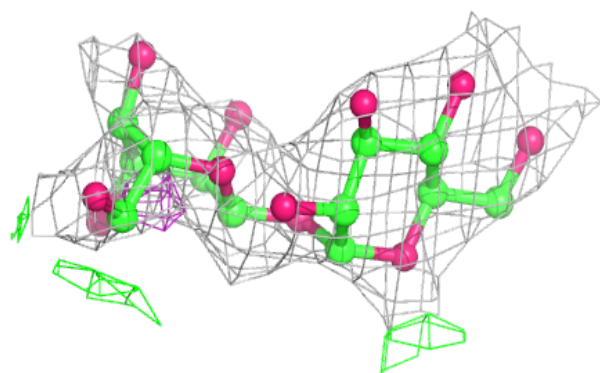
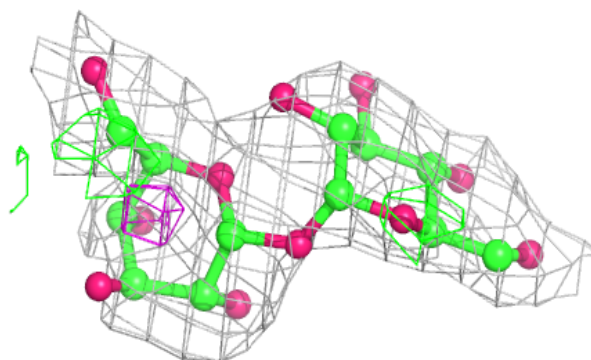
**Electron density around Chain K:**

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

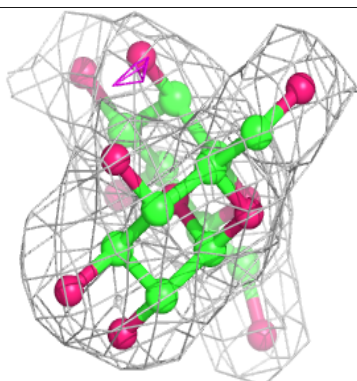
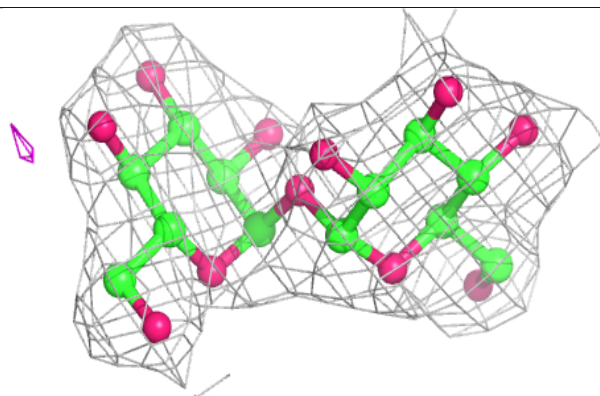
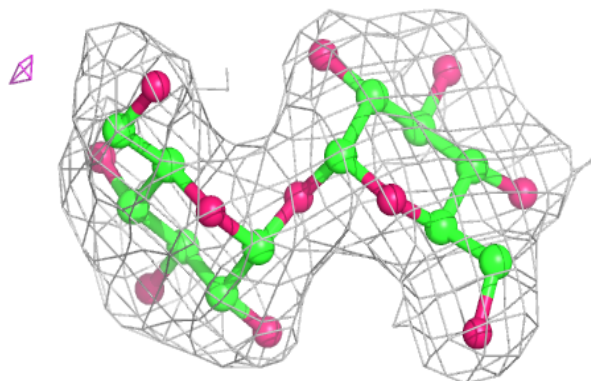


**Electron density around Chain L:**

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

**Electron density around Chain M:**

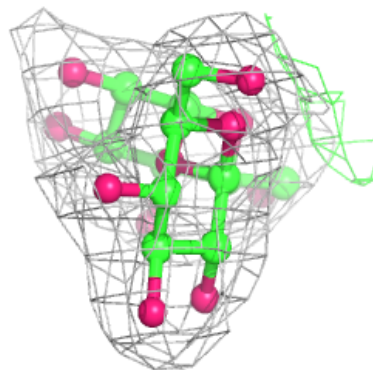
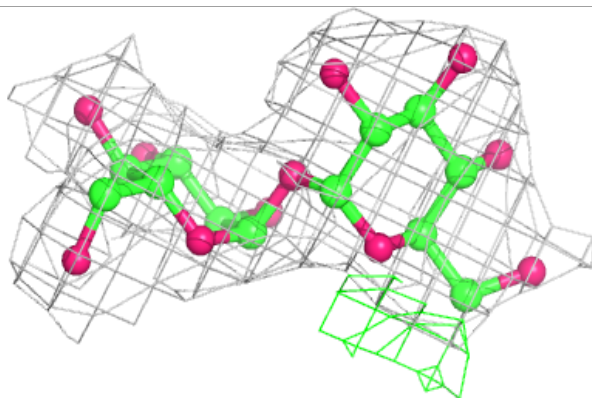
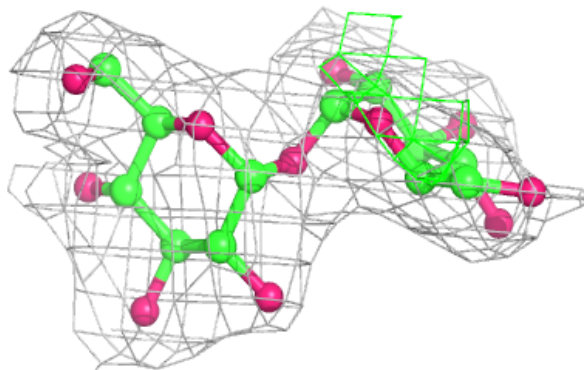
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around Chain N:**

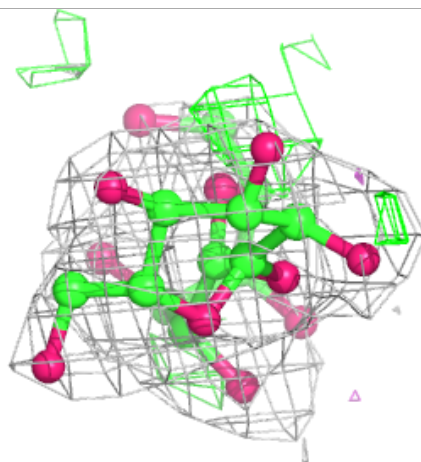
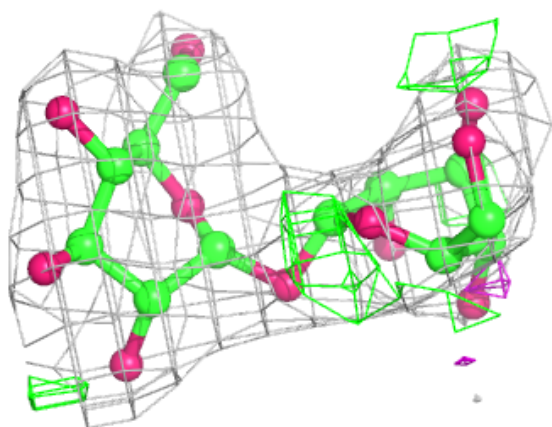
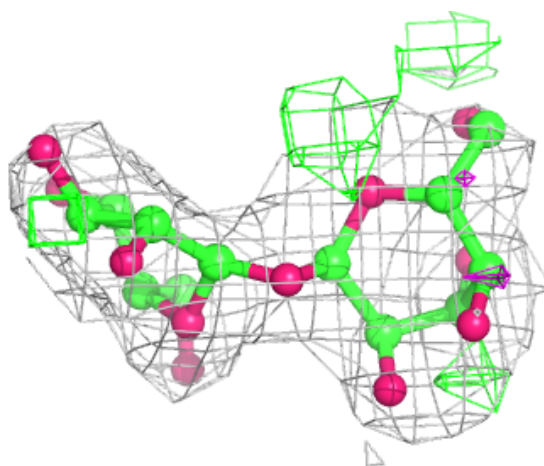
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





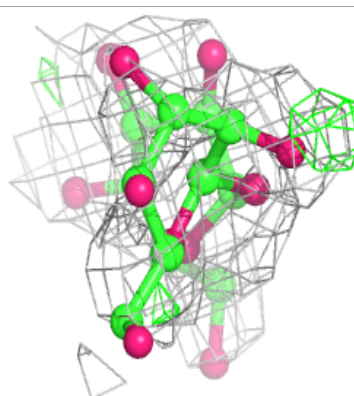
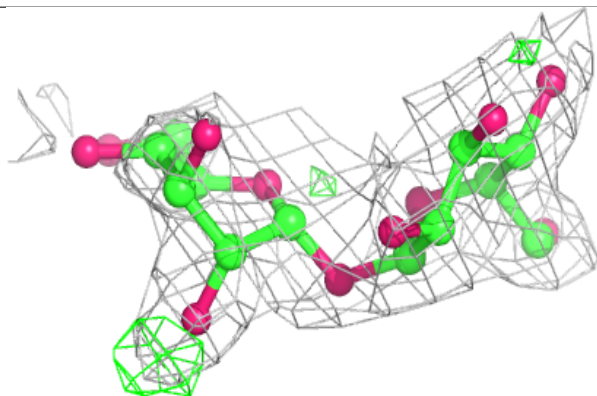
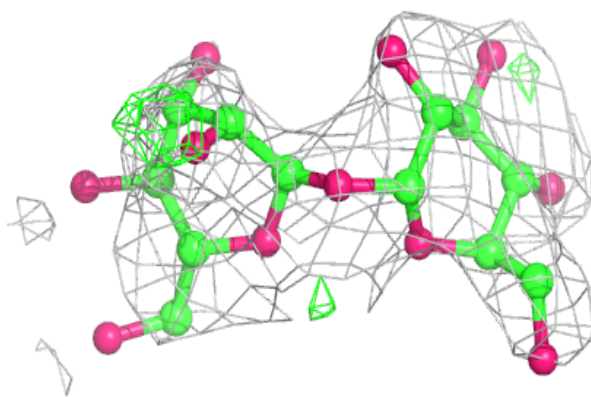
**Electron density around Chain O:**

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

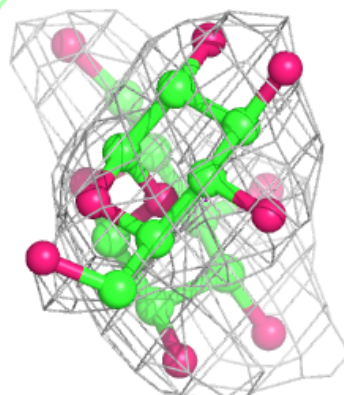
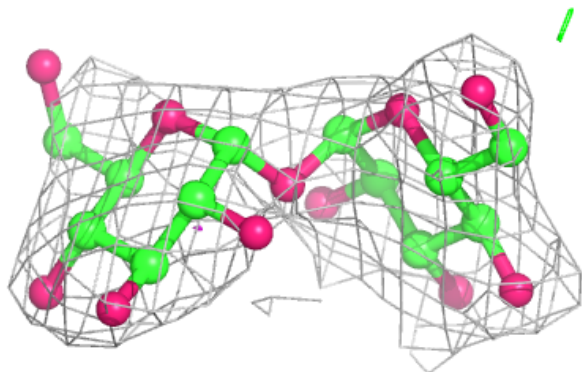
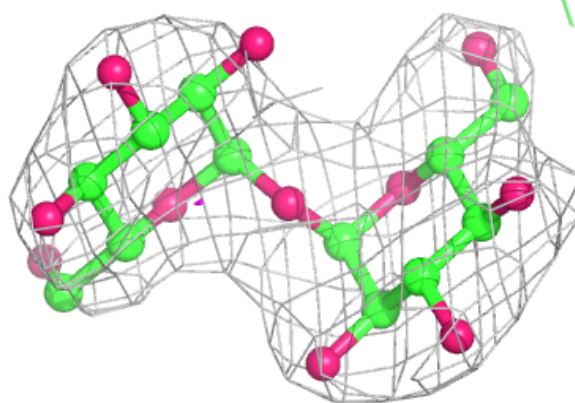


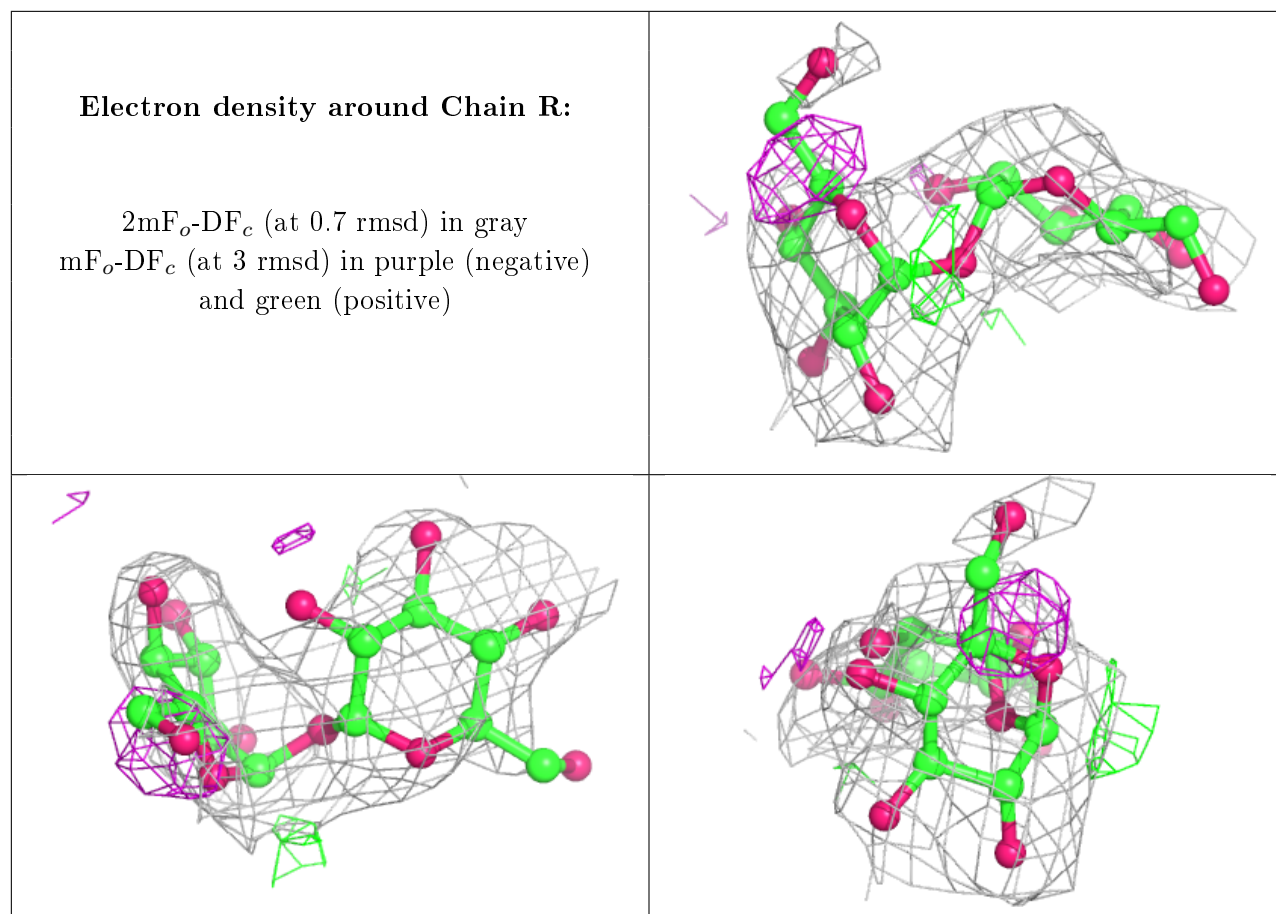
**Electron density around Chain P:**

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

**Electron density around Chain Q:**

$2mF_o-DF_c$  (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	BCN	A	613	11/11	0.69	0.27	72,91,103,104	0
5	BCN	E	609	11/11	0.69	0.31	87,98,104,106	0
5	BCN	A	612	11/11	0.73	0.25	68,74,80,84	0
5	BCN	E	612	11/11	0.76	0.33	65,92,99,102	0
5	BCN	A	614	11/11	0.78	0.26	75,86,92,95	0
5	BCN	A	615	11/11	0.80	0.25	75,96,105,107	0
5	BCN	E	611	11/11	0.83	0.22	66,83,90,92	0
5	BCN	E	610	11/11	0.85	0.18	67,83,100,101	0
5	BCN	A	611	11/11	0.89	0.21	47,60,73,73	0
4	ZN	A	602	1/1	0.94	0.07	83,83,83,83	0
4	ZN	A	601	1/1	0.98	0.17	45,45,45,45	0
4	ZN	E	601	1/1	0.98	0.06	73,73,73,73	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	ZN	E	602	1/1	0.98	0.19	48,48,48,48	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.