



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 2, 2023 – 01:24 pm GMT

PDB ID : 1W2T  
Title : beta-fructosidase from *Thermotoga maritima* in complex with raffinose  
Authors : Alberto, F.; Henrissat, B.; Czjzek, M.  
Deposited on : 2004-07-08  
Resolution : 1.87 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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

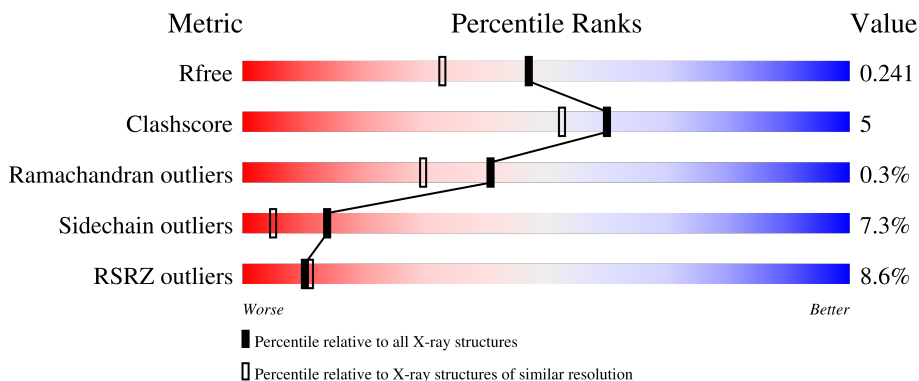
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.87 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.





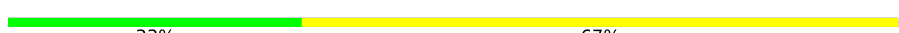
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

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	432	 5% 86% 13%
1	B	432	 11% 82% 16%
1	C	432	 9% 85% 13%
1	D	432	 8% 84% 14%
1	E	432	 10% 88% 12%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	432	 8% 87% 12%
2	G	3	 33% 67%
2	H	3	 33% 67%
2	I	3	 33% 67%
2	J	3	 33% 67%
2	K	3	 33% 67%
2	L	3	 33% 67%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 23234 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called BETA FRUCTOSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	432	3517	2250	589	665	13	0	0	0
1	B	432	3519	2252	589	665	13	0	0	0
1	C	432	3519	2252	589	665	13	0	0	0
1	D	432	3518	2251	589	665	13	0	0	0
1	E	432	3518	2251	589	665	13	0	0	0
1	F	432	3518	2252	588	665	13	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	190	ASP	GLU	engineered mutation	UNP O33833
A	108	VAL	ALA	cloning artifact	UNP O33833
A	179	ALA	VAL	cloning artifact	UNP O33833
B	190	ASP	GLU	engineered mutation	UNP O33833
B	108	VAL	ALA	cloning artifact	UNP O33833
B	179	ALA	VAL	cloning artifact	UNP O33833
C	190	ASP	GLU	engineered mutation	UNP O33833
C	108	VAL	ALA	cloning artifact	UNP O33833
C	179	ALA	VAL	cloning artifact	UNP O33833
D	190	ASP	GLU	engineered mutation	UNP O33833
D	108	VAL	ALA	cloning artifact	UNP O33833
D	179	ALA	VAL	cloning artifact	UNP O33833
E	190	ASP	GLU	engineered mutation	UNP O33833
E	108	VAL	ALA	cloning artifact	UNP O33833
E	179	ALA	VAL	cloning artifact	UNP O33833
F	190	ASP	GLU	engineered mutation	UNP O33833
F	108	VAL	ALA	cloning artifact	UNP O33833

*Continued on next page...*

Continued from previous page...

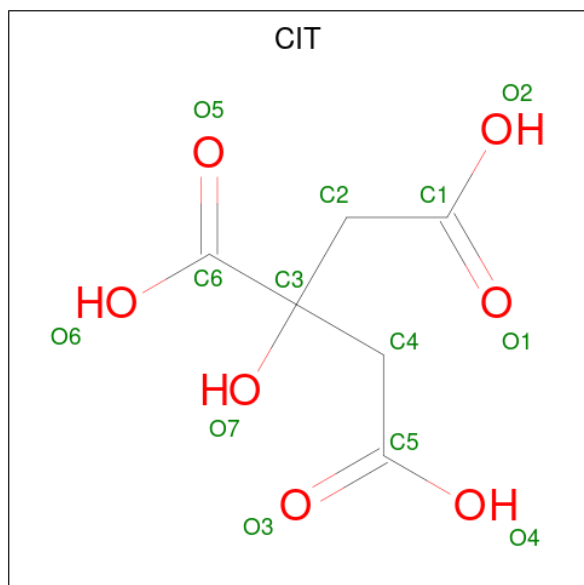
Chain	Residue	Modelled	Actual	Comment	Reference
F	179	ALA	VAL	cloning artifact	UNP O33833

- Molecule 2 is an oligosaccharide called beta-D-fructofuranose-(2-1)-[alpha-D-galactopyranoside-(1-6)]alpha-D-glucopyranose.



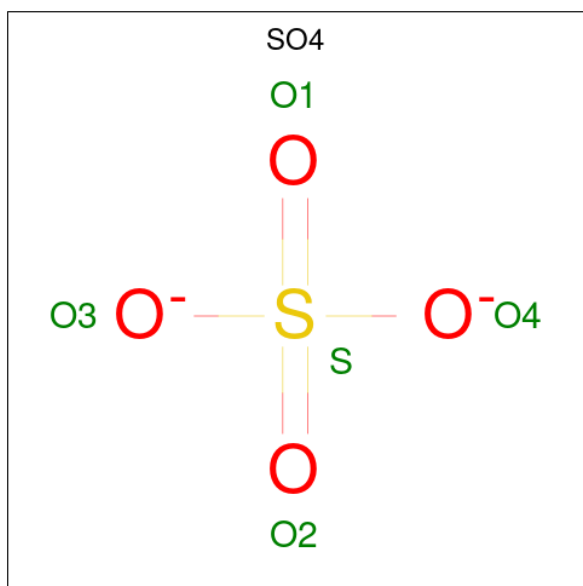
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	G	3	Total	C	O	0	0	0
			34	18	16			
2	H	3	Total	C	O	0	0	0
			34	18	16			
2	I	3	Total	C	O	0	0	0
			34	18	16			
2	J	3	Total	C	O	0	0	0
			34	18	16			
2	K	3	Total	C	O	0	0	0
			34	18	16			
2	L	3	Total	C	O	0	0	0
			34	18	16			

- Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula: C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			13	6	7		

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



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	339	Total	O	0	0
			339	339		
5	B	259	Total	O	0	0
			259	259		
5	C	316	Total	O	0	0
			316	316		
5	D	364	Total	O	0	0
			364	364		
5	E	296	Total	O	0	0
			296	296		

*Continued on next page...*

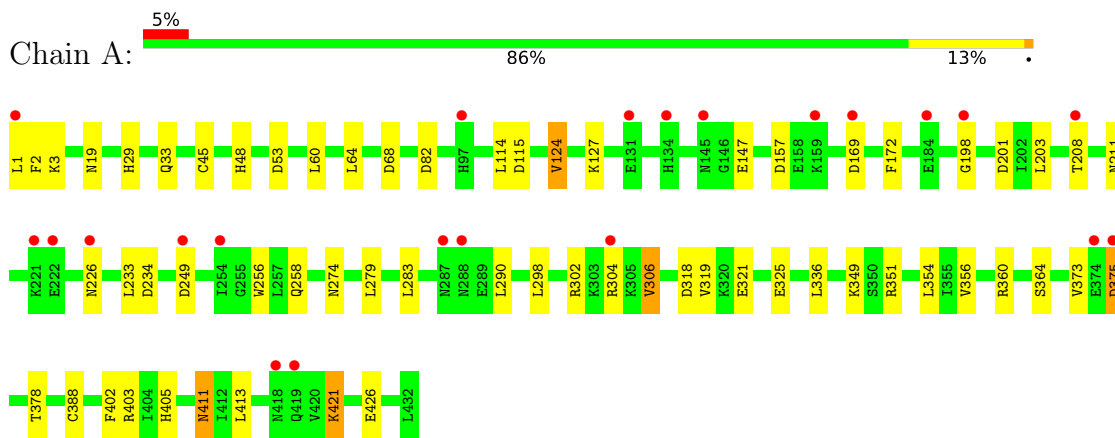
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	F	314	Total 314	O 314	0	0

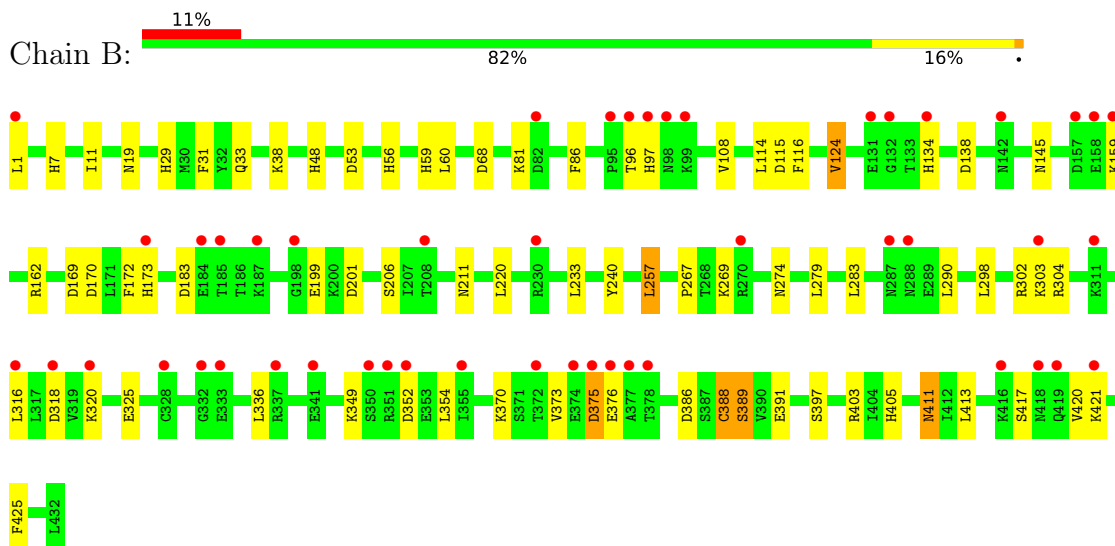
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

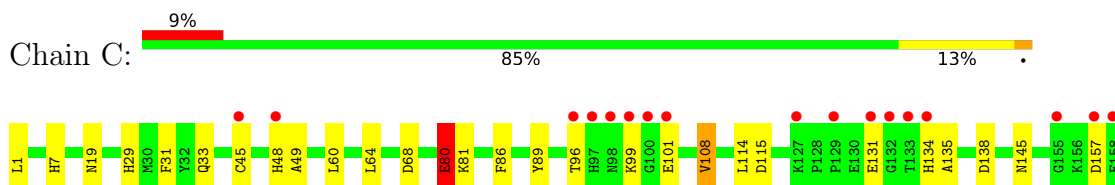
#### • Molecule 1: BETA FRUCTOSIDASE



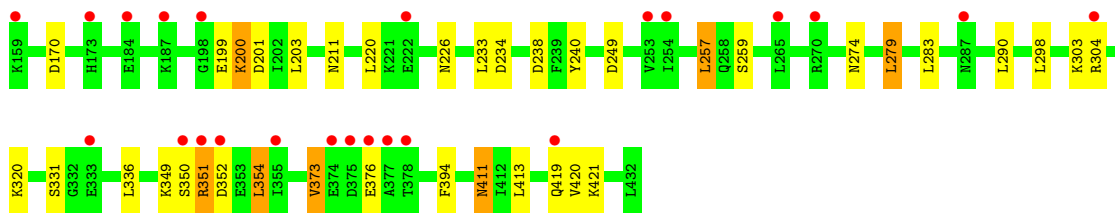
#### • Molecule 1: BETA FRUCTOSIDASE



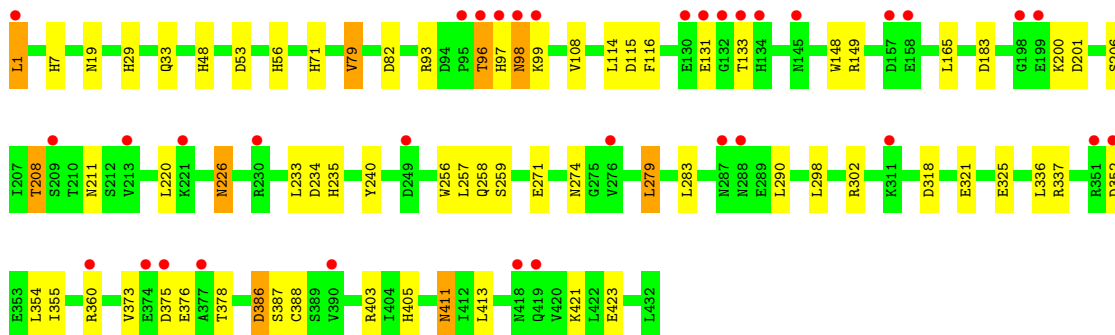
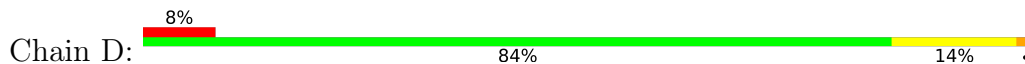
#### • Molecule 1: BETA FRUCTOSIDASE



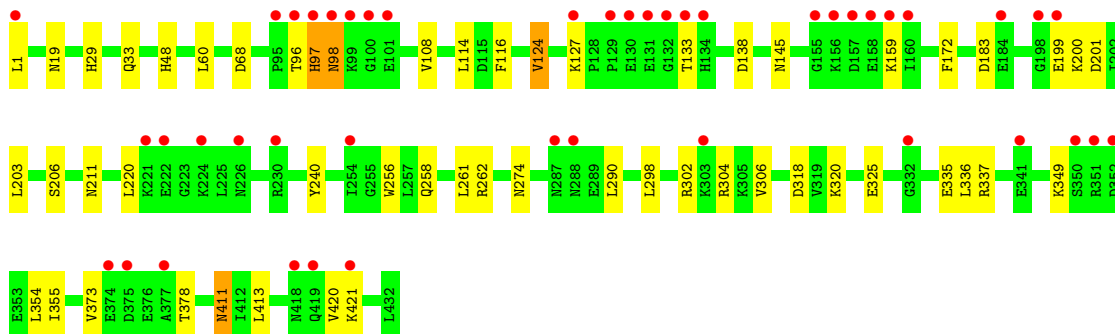
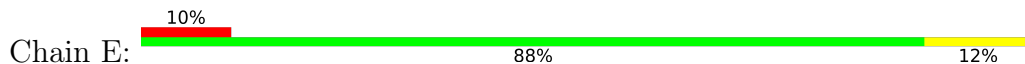




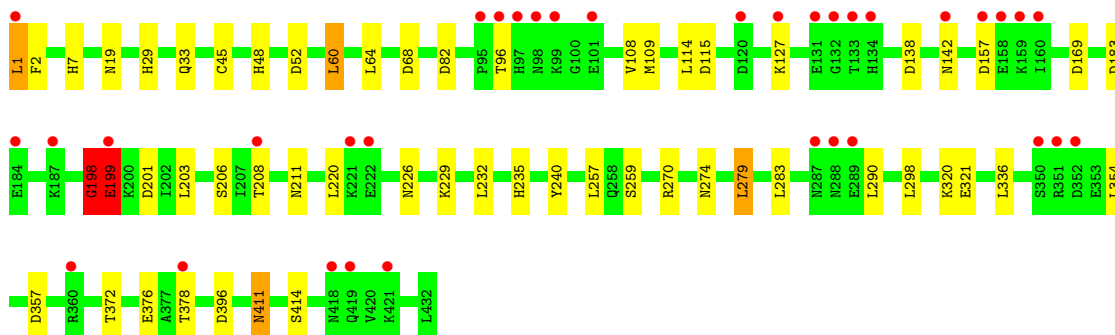
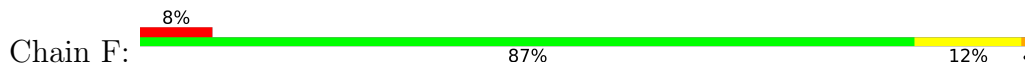
● Molecule 1: BETA FRUCTOSIDASE



● Molecule 1: BETA FRUCTOSIDASE



● Molecule 1: BETA FRUCTOSIDASE

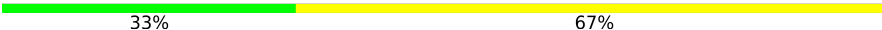


- Molecule 2: beta-D-fructofuranose-(2-1)-[alpha-D-galactopyranose-(1-6)]alpha-D-glucopyranose

Chain G:  33% 67%

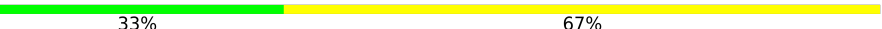
GLC1  
FRU2  
GLA3

- Molecule 2: beta-D-fructofuranose-(2-1)-[alpha-D-galactopyranose-(1-6)]alpha-D-glucopyranose

Chain H:  33% 67%

GLC1  
FRU2  
GLA3

- Molecule 2: beta-D-fructofuranose-(2-1)-[alpha-D-galactopyranose-(1-6)]alpha-D-glucopyranose

Chain I:  33% 67%

GLC1  
FRU2  
GLA3

- Molecule 2: beta-D-fructofuranose-(2-1)-[alpha-D-galactopyranose-(1-6)]alpha-D-glucopyranose

Chain J:  33% 67%

GLC1  
FRU2  
GLA3

- Molecule 2: beta-D-fructofuranose-(2-1)-[alpha-D-galactopyranose-(1-6)]alpha-D-glucopyranose

Chain K:  33% 67%

GLC1  
FRU2  
GLA3

- Molecule 2: beta-D-fructofuranose-(2-1)-[alpha-D-galactopyranose-(1-6)]alpha-D-glucopyranose

Chain L:  33% 67%

GLC1  
FRU2  
GLA3

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.50Å 114.70Å 130.00Å 90.00° 98.90° 90.00°	Depositor
Resolution (Å)	40.00 – 1.87 32.11 – 1.87	Depositor EDS
% Data completeness (in resolution range)	97.7 (40.00-1.87) 97.7 (32.11-1.87)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.97 (at 1.87Å)	Xtrriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.198 , 0.229 0.213 , 0.241	Depositor DCC
$R_{free}$ test set	11055 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.7	Xtrriage
Anisotropy	0.198	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 52.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	23234	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.89% 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: GLA, GLC, CIT, SO4, FRU

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	0.34	1/3607 (0.0%)	0.73	13/4881 (0.3%)
1	B	0.32	0/3610	0.71	8/4885 (0.2%)
1	C	0.39	1/3610 (0.0%)	0.70	9/4885 (0.2%)
1	D	0.32	0/3608	0.70	8/4883 (0.2%)
1	E	0.31	0/3609	0.70	4/4883 (0.1%)
1	F	0.33	0/3609	0.74	14/4884 (0.3%)
All	All	0.34	2/21653 (0.0%)	0.71	56/29301 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1
1	C	0	1
1	F	0	2
All	All	0	4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	80	GLU	C-N	12.95	1.63	1.34
1	A	2	PHE	C-N	-5.36	1.21	1.34

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	198	GLY	CA-C-N	8.66	136.25	117.20
1	F	198	GLY	CA-C-O	-8.53	105.24	120.60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	198	GLY	O-C-N	-7.91	110.04	122.70
1	A	2	PHE	O-C-N	-7.81	110.21	122.70
1	A	1	LEU	CB-CA-C	-5.99	98.82	110.20

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	388	CYS	Mainchain
1	C	80	GLU	Mainchain
1	F	198	GLY	Peptide,Mainchain

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

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3517	0	3398	31	0
1	B	3519	0	3411	33	0
1	C	3519	0	3411	42	0
1	D	3518	0	3404	35	0
1	E	3518	0	3407	31	0
1	F	3518	0	3406	20	0
2	G	34	0	29	0	0
2	H	34	0	29	0	0
2	I	34	0	29	0	0
2	J	34	0	29	0	0
2	K	34	0	29	0	0
2	L	34	0	29	0	0
3	A	13	0	5	5	0
4	B	5	0	0	0	0
4	C	10	0	0	0	0
4	E	5	0	0	0	0
5	A	339	0	0	16	0
5	B	259	0	0	1	0
5	C	316	0	0	5	0
5	D	364	0	0	4	0
5	E	296	0	0	4	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	F	314	0	0	6	0
All	All	23234	0	20616	191	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 5.

The worst 5 of 191 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:211:ASN:OD1	5:A:2179:HOH:O	1.56	1.20
1:F:270:ARG:CZ	5:F:2203:HOH:O	1.88	1.18
1:B:267:PRO:HD2	1:B:405:HIS:CD2	1.87	1.08
3:A:1433:CIT:O2	5:A:2332:HOH:O	1.75	1.03
1:C:211:ASN:HD21	1:C:259:SER:HA	1.19	1.02

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

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

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	430/432 (100%)	411 (96%)	18 (4%)	1 (0%)	47	37
1	B	430/432 (100%)	412 (96%)	17 (4%)	1 (0%)	47	37
1	C	430/432 (100%)	414 (96%)	15 (4%)	1 (0%)	47	37
1	D	430/432 (100%)	414 (96%)	15 (4%)	1 (0%)	47	37
1	E	430/432 (100%)	414 (96%)	15 (4%)	1 (0%)	47	37
1	F	430/432 (100%)	412 (96%)	16 (4%)	2 (0%)	29	17
All	All	2580/2592 (100%)	2477 (96%)	96 (4%)	7 (0%)	41	30

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	351	ARG
1	F	199	GLU
1	D	98	ASN
1	E	98	ASN
1	B	389	SER

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	385/387 (100%)	363 (94%)	22 (6%)	20 9
1	B	387/387 (100%)	352 (91%)	35 (9%)	9 3
1	C	387/387 (100%)	359 (93%)	28 (7%)	14 5
1	D	386/387 (100%)	358 (93%)	28 (7%)	14 5
1	E	387/387 (100%)	362 (94%)	25 (6%)	17 7
1	F	386/387 (100%)	355 (92%)	31 (8%)	12 4
All	All	2318/2322 (100%)	2149 (93%)	169 (7%)	14 5

5 of 169 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	97	HIS
1	F	114	LEU
1	E	133	THR
1	E	349	LYS
1	F	211	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	7	HIS
1	F	235	HIS
1	D	235	HIS
1	F	59	HIS

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
1	F	411	ASN

### 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](#)

18 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	GLC	G	1	2	11,11,12	0.82	0	15,15,17	1.56	2 (13%)
2	FRU	G	2	2	11,12,12	2.00	1 (9%)	10,18,18	0.70	0
2	GLA	G	3	2	11,11,12	0.63	0	15,15,17	0.59	0
2	GLC	H	1	2	11,11,12	0.67	0	15,15,17	0.86	0
2	FRU	H	2	2	11,12,12	2.02	1 (9%)	10,18,18	0.47	0
2	GLA	H	3	2	11,11,12	0.55	0	15,15,17	0.74	1 (6%)
2	GLC	I	1	2	11,11,12	0.63	0	15,15,17	0.87	1 (6%)
2	FRU	I	2	2	11,12,12	2.04	1 (9%)	10,18,18	0.61	0
2	GLA	I	3	2	11,11,12	0.58	0	15,15,17	0.62	0
2	GLC	J	1	2	11,11,12	0.99	1 (9%)	15,15,17	2.13	3 (20%)
2	FRU	J	2	2	11,12,12	1.95	1 (9%)	10,18,18	0.56	0
2	GLA	J	3	2	11,11,12	0.61	0	15,15,17	0.55	0
2	GLC	K	1	2	11,11,12	0.83	0	15,15,17	1.30	2 (13%)
2	FRU	K	2	2	11,12,12	2.01	1 (9%)	10,18,18	0.66	0
2	GLA	K	3	2	11,11,12	0.55	0	15,15,17	0.54	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GLC	L	1	2	11,11,12	0.68	0	15,15,17	1.49	2 (13%)
2	FRU	L	2	2	11,12,12	2.02	1 (9%)	10,18,18	0.46	0
2	GLA	L	3	2	11,11,12	0.56	0	15,15,17	0.68	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
2	GLC	G	1	2	-	0/2/19/22	0/1/1/1
2	FRU	G	2	2	-	3/5/24/24	0/1/1/1
2	GLA	G	3	2	-	2/2/19/22	0/1/1/1
2	GLC	H	1	2	-	0/2/19/22	0/1/1/1
2	FRU	H	2	2	-	3/5/24/24	0/1/1/1
2	GLA	H	3	2	-	2/2/19/22	0/1/1/1
2	GLC	I	1	2	-	0/2/19/22	0/1/1/1
2	FRU	I	2	2	-	3/5/24/24	0/1/1/1
2	GLA	I	3	2	-	1/2/19/22	0/1/1/1
2	GLC	J	1	2	-	0/2/19/22	0/1/1/1
2	FRU	J	2	2	-	3/5/24/24	0/1/1/1
2	GLA	J	3	2	-	2/2/19/22	0/1/1/1
2	GLC	K	1	2	-	0/2/19/22	0/1/1/1
2	FRU	K	2	2	-	3/5/24/24	0/1/1/1
2	GLA	K	3	2	-	0/2/19/22	0/1/1/1
2	GLC	L	1	2	-	0/2/19/22	0/1/1/1
2	FRU	L	2	2	-	3/5/24/24	0/1/1/1
2	GLA	L	3	2	-	1/2/19/22	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	K	2	FRU	O6-C6	-6.52	1.14	1.42
2	I	2	FRU	O6-C6	-6.49	1.15	1.42
2	H	2	FRU	O6-C6	-6.42	1.15	1.42
2	G	2	FRU	O6-C6	-6.40	1.15	1.42
2	L	2	FRU	O6-C6	-6.37	1.15	1.42

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	J	1	GLC	C1-O5-C5	-4.70	105.82	112.19
2	L	1	GLC	O5-C1-C2	-4.53	103.77	110.77
2	J	1	GLC	C1-C2-C3	-4.11	104.62	109.67
2	J	1	GLC	O5-C1-C2	4.10	117.11	110.77
2	G	1	GLC	C1-O5-C5	-3.71	107.17	112.19

There are no chirality outliers.

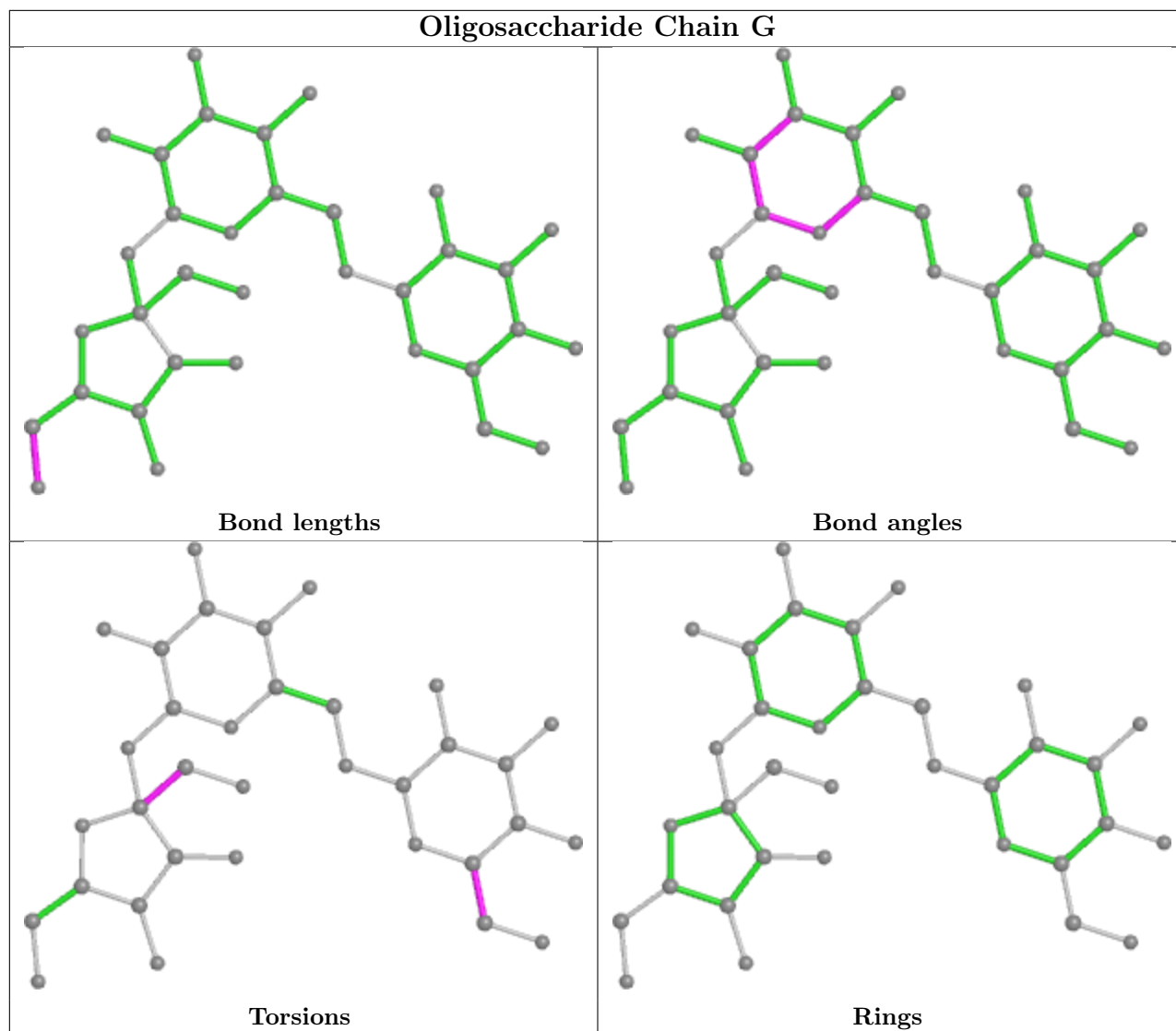
5 of 26 torsion outliers are listed below:

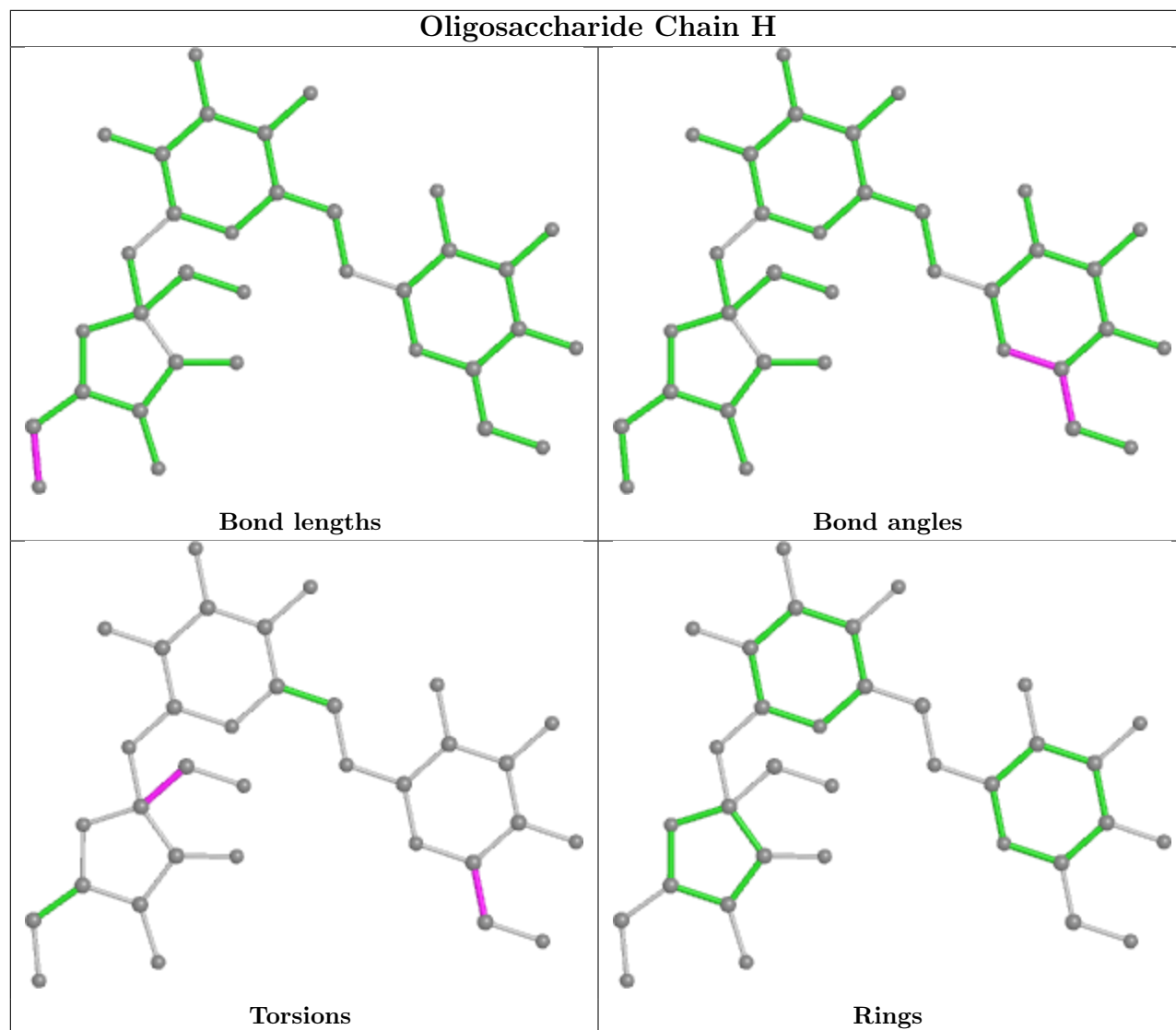
Mol	Chain	Res	Type	Atoms
2	G	2	FRU	O1-C1-C2-O2
2	H	2	FRU	O1-C1-C2-O2
2	I	2	FRU	O1-C1-C2-O2
2	J	2	FRU	O1-C1-C2-O2
2	K	2	FRU	O1-C1-C2-O2

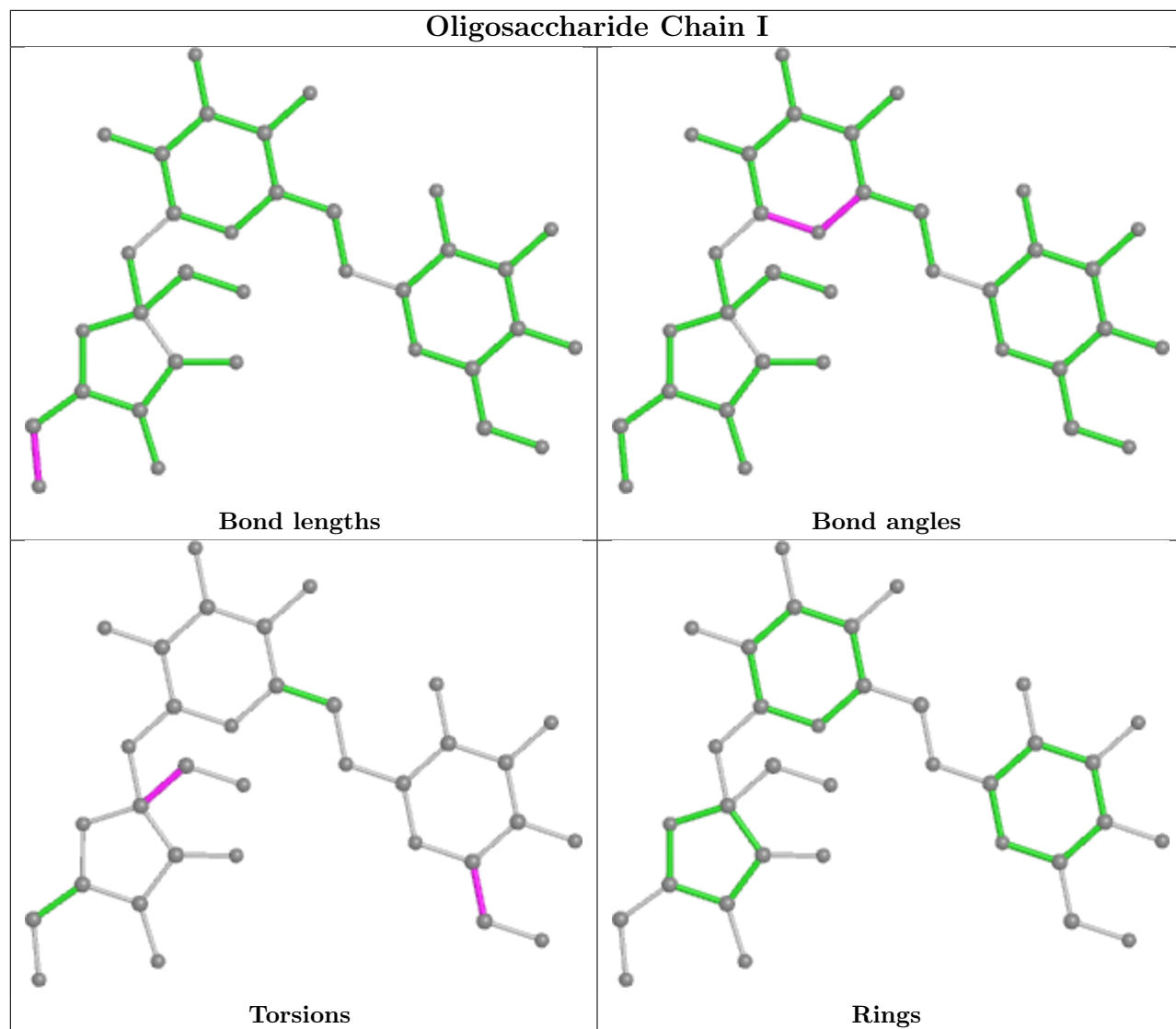
There are no ring outliers.

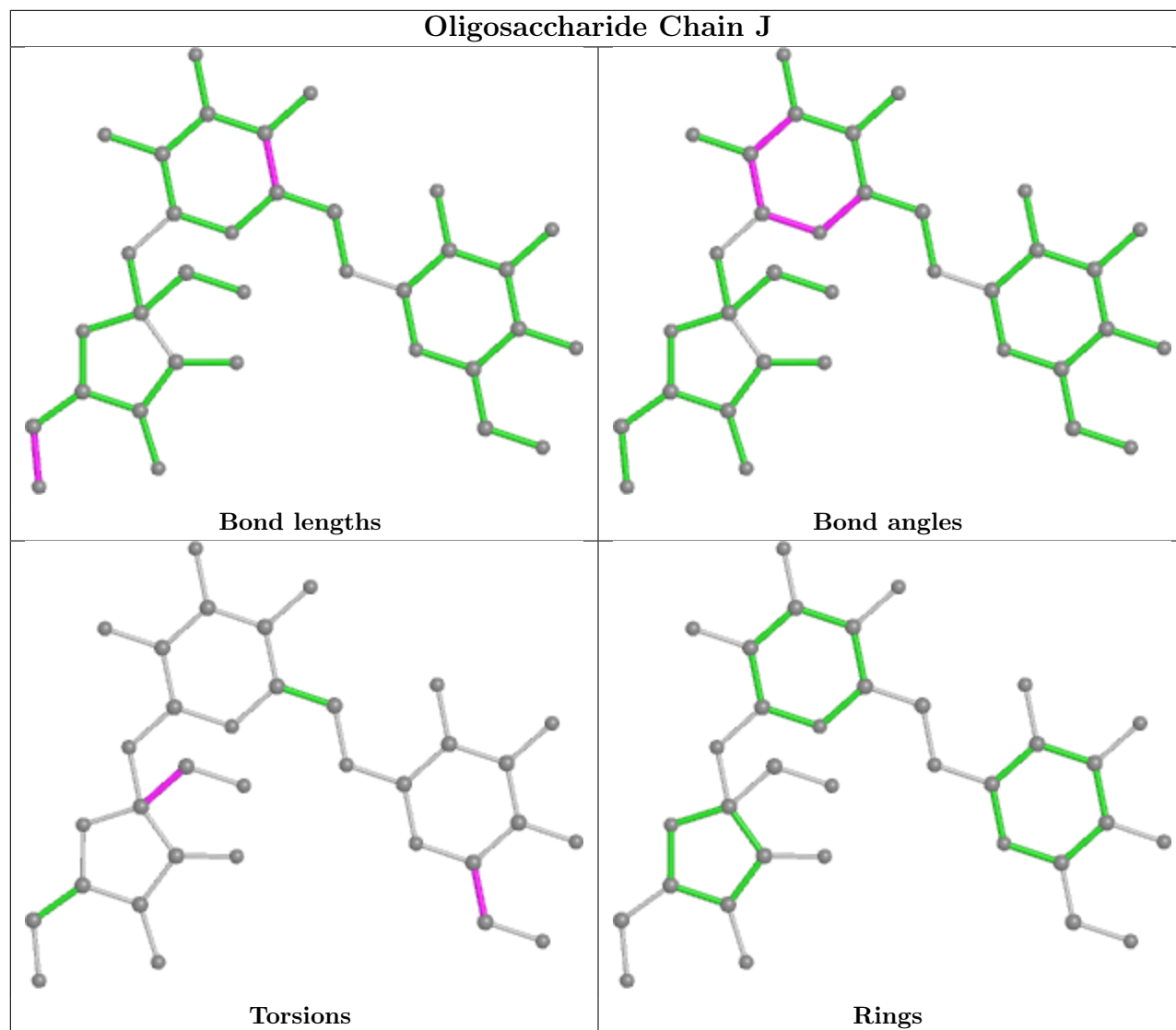
No monomer is involved in short contacts.

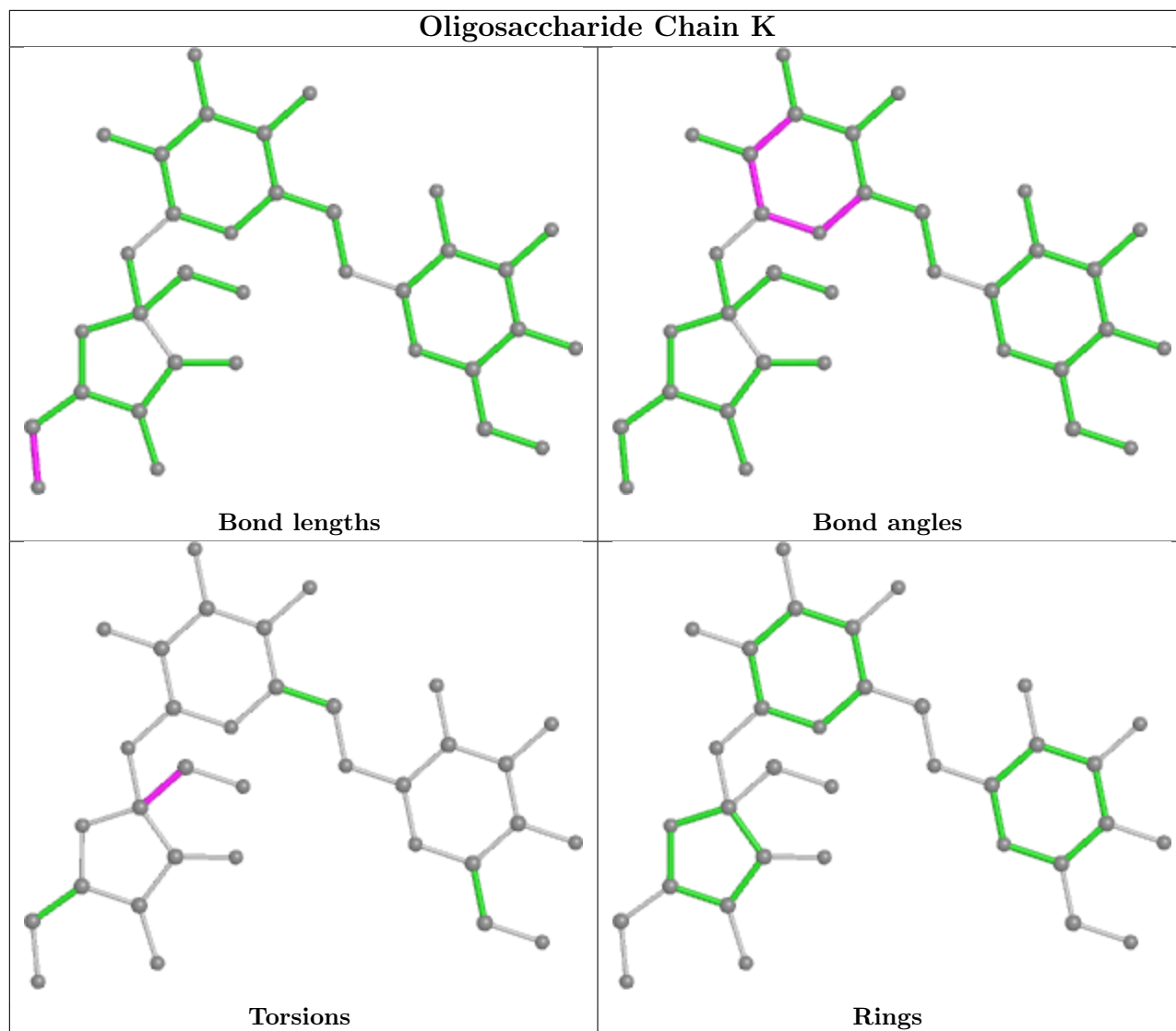
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

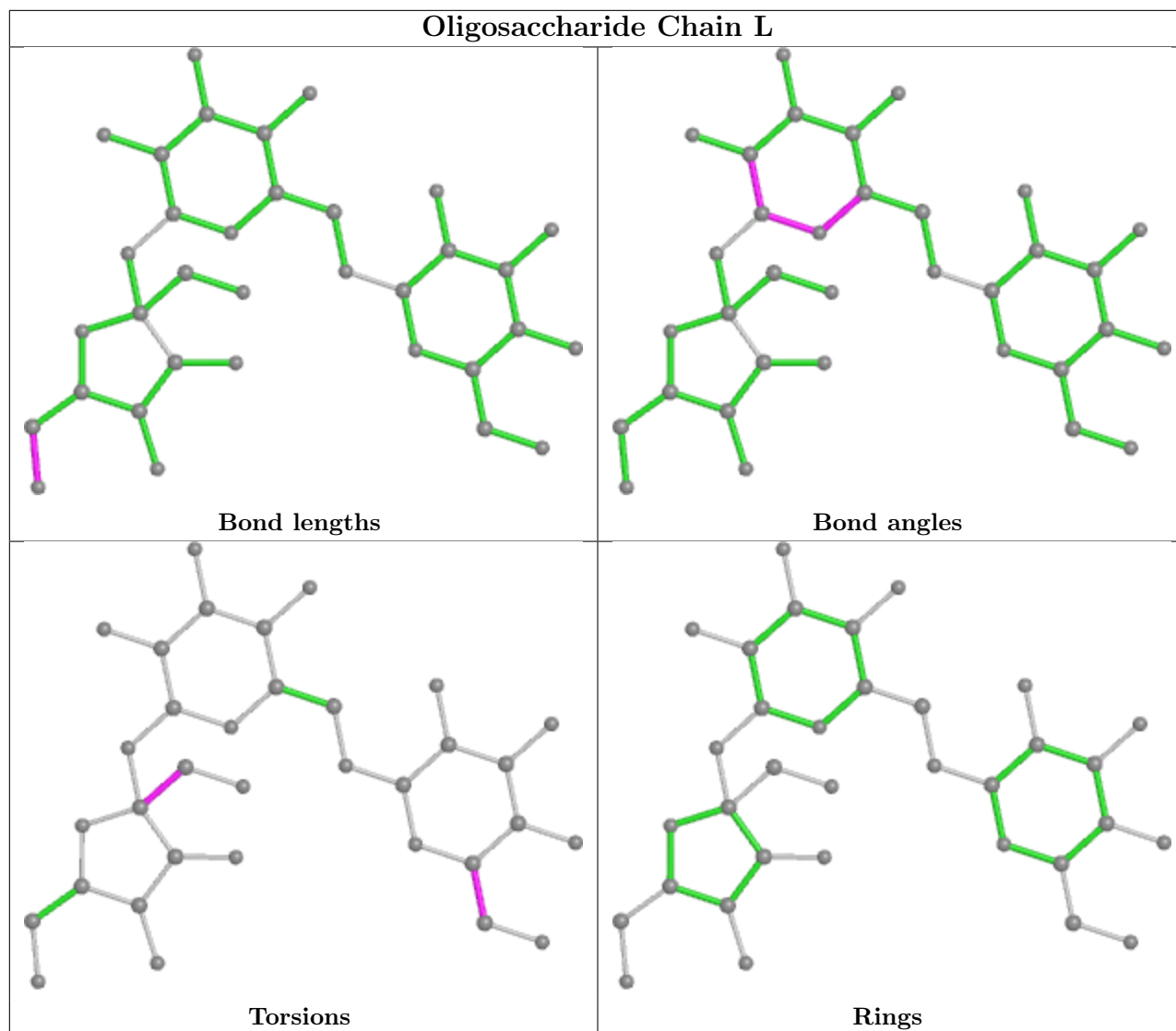












## 5.6 Ligand geometry [i](#)

5 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	SO4	B	1433	-	4,4,4	0.16	0	6,6,6	0.17	0
4	SO4	E	1433	-	4,4,4	0.55	0	6,6,6	0.88	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	CIT	A	1433	-	12,12,12	1.04	0	17,17,17	1.76	4 (23%)
4	SO4	C	1434	-	4,4,4	0.14	0	6,6,6	0.12	0
4	SO4	C	1433	-	4,4,4	0.23	0	6,6,6	0.35	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
3	CIT	A	1433	-	-	4/16/16/16	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1433	CIT	O6-C6-C3	3.63	119.35	113.05
3	A	1433	CIT	C4-C3-C2	3.18	117.45	109.16
3	A	1433	CIT	C2-C3-C6	-2.73	104.23	110.11
3	A	1433	CIT	C4-C3-C6	-2.14	105.50	110.11

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1433	CIT	C1-C2-C3-O7
3	A	1433	CIT	C1-C2-C3-C6
3	A	1433	CIT	C1-C2-C3-C4
3	A	1433	CIT	C2-C3-C4-C5

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1433	CIT	5	0

## 5.7 Other polymers [\(i\)](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	80:GLU	C	81:LYS	N	1.63

## 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	432/432 (100%)	0.32	22 (5%) 28 29	7, 16, 28, 33	4 (0%)
1	B	432/432 (100%)	0.63	48 (11%) 5 5	8, 17, 28, 34	6 (1%)
1	C	432/432 (100%)	0.50	40 (9%) 8 9	7, 17, 28, 34	8 (1%)
1	D	432/432 (100%)	0.52	34 (7%) 12 13	8, 16, 28, 34	8 (1%)
1	E	432/432 (100%)	0.50	44 (10%) 6 7	8, 16, 28, 34	9 (2%)
1	F	432/432 (100%)	0.51	35 (8%) 12 13	7, 16, 28, 34	10 (2%)
All	All	2592/2592 (100%)	0.50	223 (8%) 10 11	7, 16, 28, 34	45 (1%)

The worst 5 of 223 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	98	ASN	11.9
1	D	98	ASN	10.9
1	E	97	HIS	10.5
1	B	97	HIS	10.3
1	C	97	HIS	10.2

### 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](#)

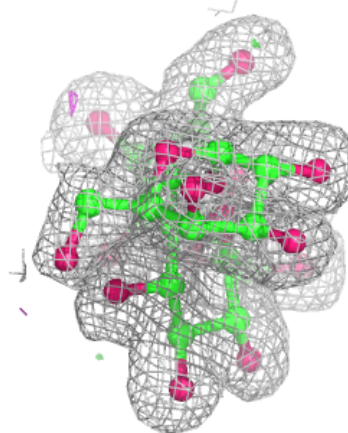
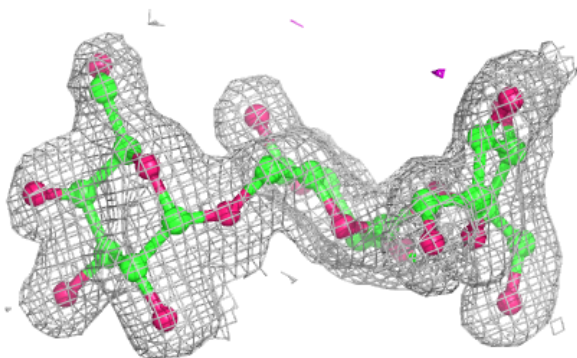
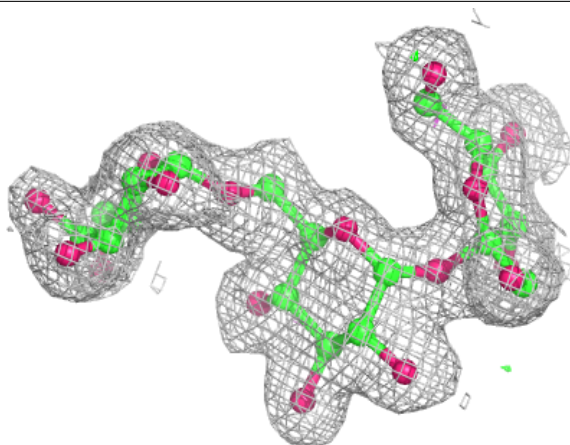
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	FRU	I	2	12/12	0.65	0.23	24,28,29,31	0
2	GLA	H	3	11/12	0.68	0.22	41,42,42,43	0
2	GLA	I	3	11/12	0.72	0.27	43,44,45,45	0
2	GLC	H	1	11/12	0.75	0.16	31,33,34,36	0
2	FRU	H	2	12/12	0.76	0.17	25,27,27,29	0
2	GLC	I	1	11/12	0.76	0.15	33,34,36,37	0
2	GLA	G	3	11/12	0.80	0.17	29,30,31,31	0
2	FRU	K	2	12/12	0.81	0.16	21,23,25,25	0
2	GLA	K	3	11/12	0.81	0.17	35,36,37,38	0
2	GLA	L	3	11/12	0.85	0.27	33,35,36,36	0
2	FRU	L	2	12/12	0.88	0.14	18,19,20,21	0
2	GLA	J	3	11/12	0.89	0.17	29,31,32,32	0
2	GLC	K	1	11/12	0.89	0.11	27,28,29,30	0
2	GLC	L	1	11/12	0.90	0.12	23,24,25,26	0
2	FRU	G	2	12/12	0.90	0.14	17,19,20,21	0
2	FRU	J	2	12/12	0.90	0.14	16,17,18,19	0
2	GLC	J	1	11/12	0.91	0.12	20,21,22,24	0
2	GLC	G	1	11/12	0.92	0.09	23,24,25,27	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.

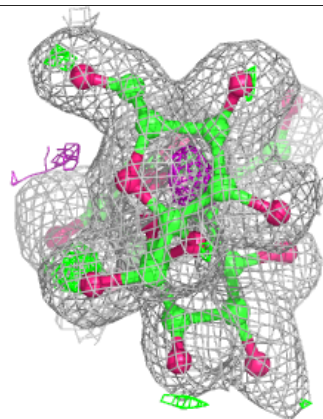
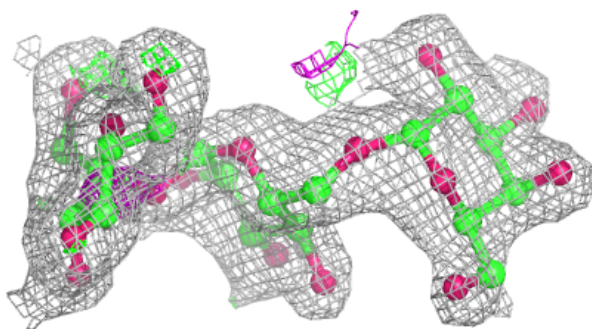
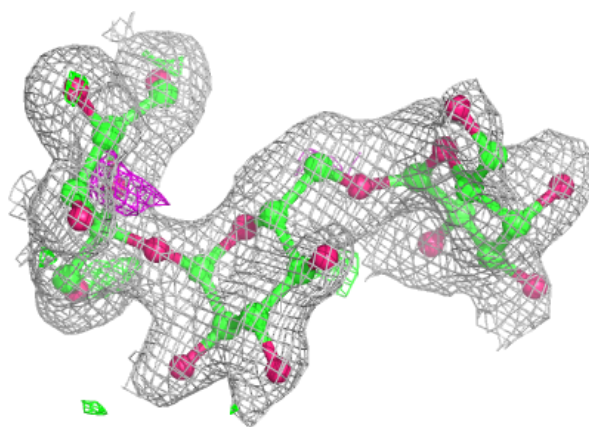
**Electron density around Chain G:**

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



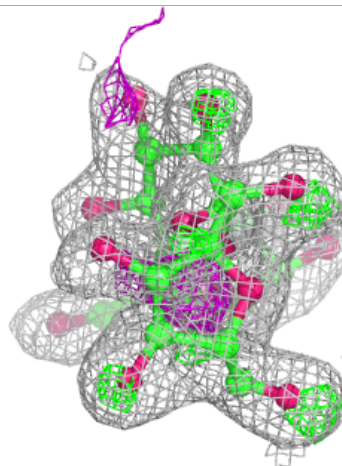
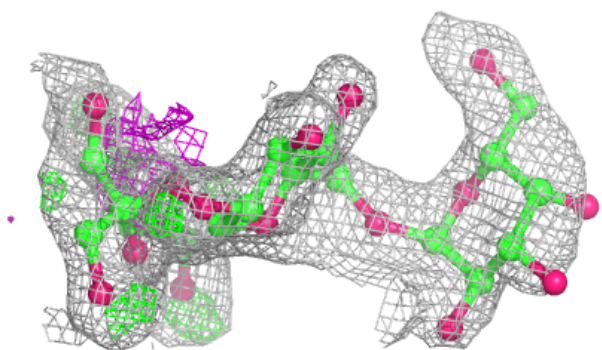
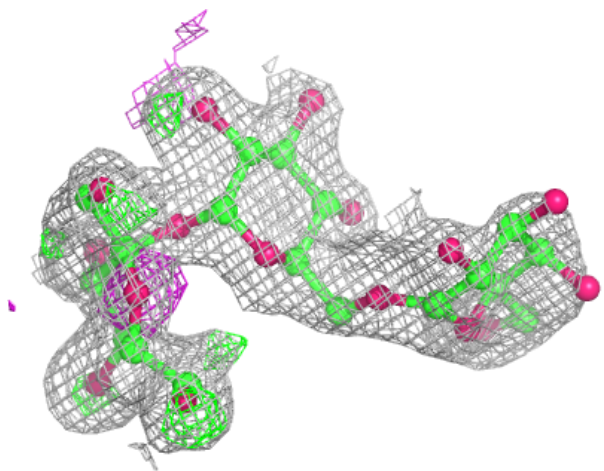
**Electron density around Chain H:**

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



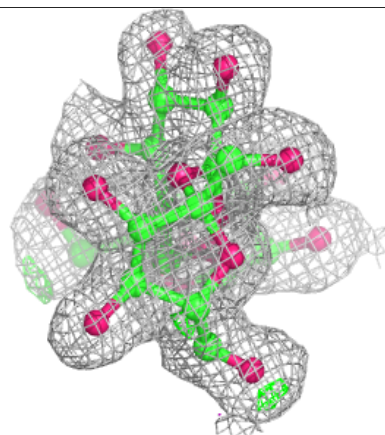
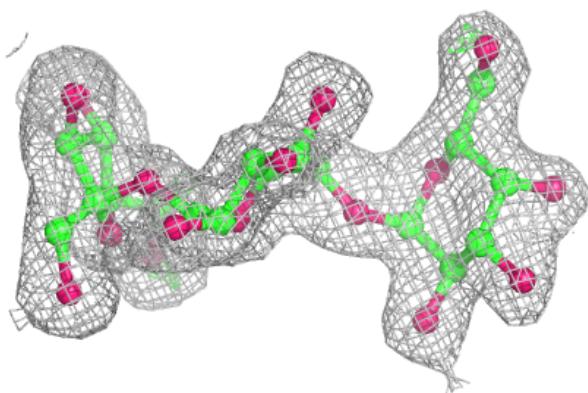
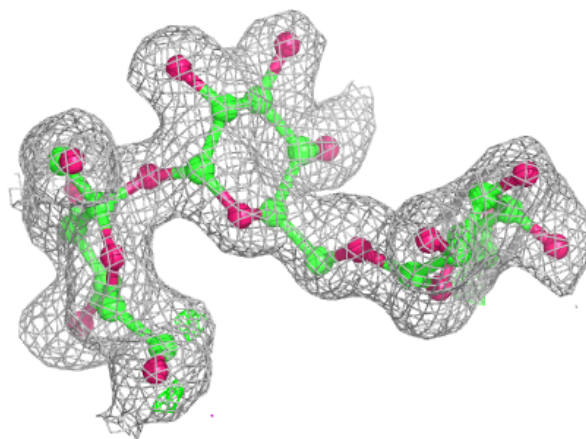
**Electron density around Chain I:**

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



**Electron density around Chain J:**

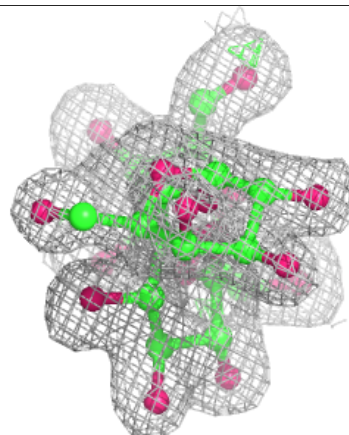
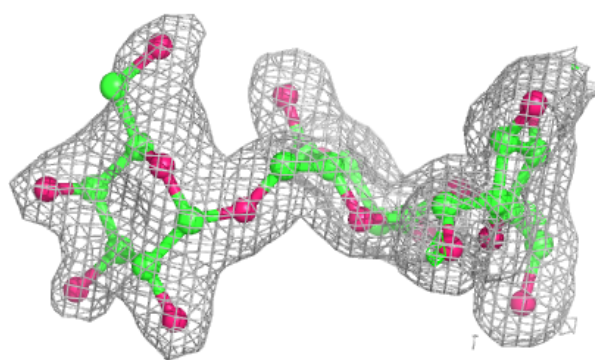
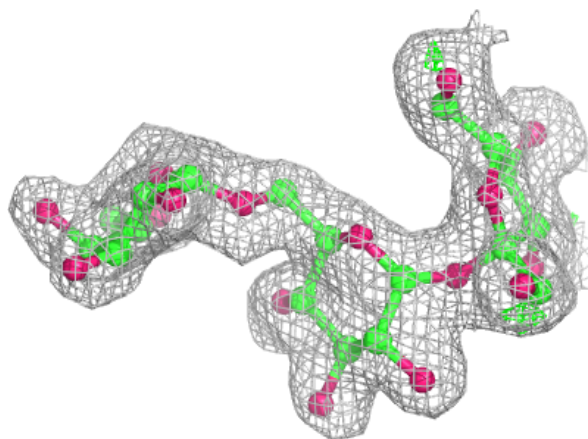
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

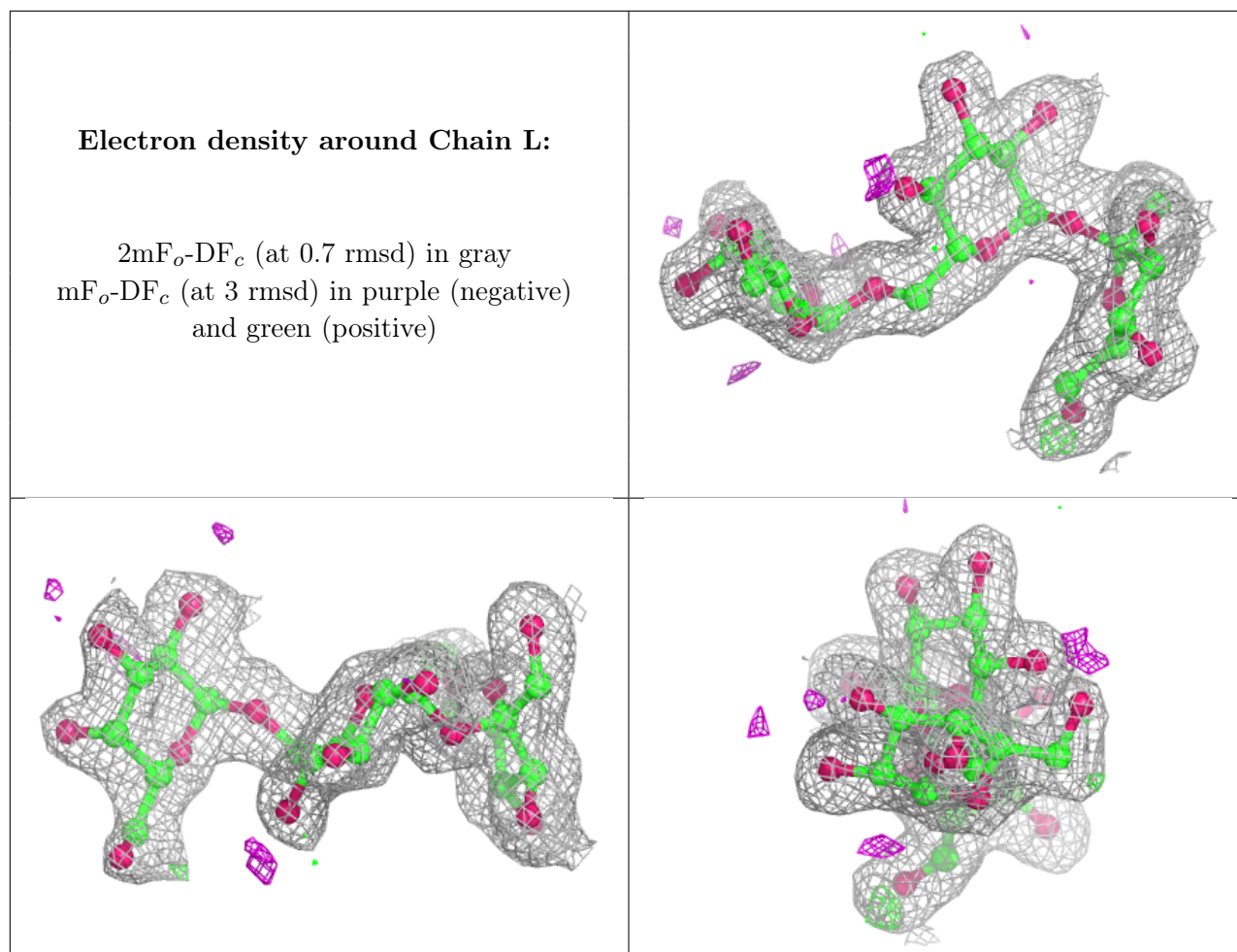




**Electron density around Chain K:**

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





## 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( $\text{\AA}^2$ )	Q<0.9
3	CIT	A	1433	13/13	0.75	0.23	50,51,51,52	0
4	SO4	C	1434	5/5	0.82	0.24	44,45,45,46	0
4	SO4	B	1433	5/5	0.85	0.19	43,44,44,45	0
4	SO4	C	1433	5/5	0.89	0.44	2,2,13,14	0
4	SO4	E	1433	5/5	0.94	0.39	2,2,2,2	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.