



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2025 – 10:20 PM JST

PDB ID : 9KYR / pdb\_00009kyr  
Title : GH57 family amylopullulanase from Aquifex aeolicus D352N mutant complex with gama-cyclodextrin  
Authors : Zhu, Z.M.; Wang, W.W.; Yu, F.; Li, M.J.; Xu, Q.; Zhou, H.; Huang, L.Q.; Wang, Q.S.  
Deposited on : 2024-12-09  
Resolution : 1.89 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:

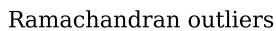
MolProbity : 4-5-2 with Phenix2.0rc1  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 2.0rc1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.006 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.43.1

**i**

## X-RAY DIFFRACTION

A.

Metric	Percentile Banks	Value
--------	------------------	-------



R <sub>free</sub>
Clashscore
Ramachandran outliers
Sidechain outliers
RSRZ outliers

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.

1	A	477	<div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div></div> <div>83%14%</div>
1	B	477	<div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div></div> <div>81%18%</div>
1	C	477	<div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div></div> <div>82%15%</div>
1	D	477	<div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div></div> <div>90%9%</div>
2	E	8	<div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div></div> <div>38%62%</div>
2	F	8	<div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div></div> <div>25%25%50%</div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	8	
2	H	8	

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
2	GLC	E	2	-	-	X	-
5	EDO	B	504	-	-	X	-
5	EDO	B	505	-	-	X	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 17203 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.

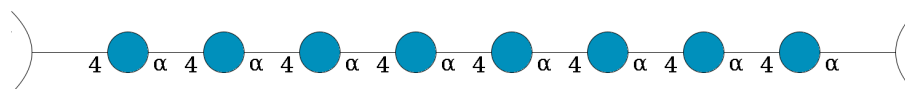
- Molecule 1 is a protein called Glycoside hydrolase family 57 N-terminal domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	476	Total	C	N	O	S	0	2	0
			4034	2639	647	738	10			
1	B	476	Total	C	N	O	S	0	0	0
			4020	2631	645	735	9			
1	C	476	Total	C	N	O	S	0	1	0
			4026	2634	646	737	9			
1	D	477	Total	C	N	O	S	0	1	0
			4034	2639	647	738	10			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	352	ASN	ASP	engineered mutation	UNP O66934
B	352	ASN	ASP	engineered mutation	UNP O66934
C	352	ASN	ASP	engineered mutation	UNP O66934
D	352	ASN	ASP	engineered mutation	UNP O66934

- Molecule 2 is an oligosaccharide called Cyclooctakis-(1-4)-(alpha-D-glucopyranose).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	E	8	Total	C	O	0	0	0
			88	48	40			
2	F	8	Total	C	O	0	0	0
			88	48	40			
2	G	8	Total	C	O	0	0	0
			88	48	40			

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	H	8	Total	C	O	0	0	0
			88	48	40			

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula:  $C_3H_8O_3$ ).



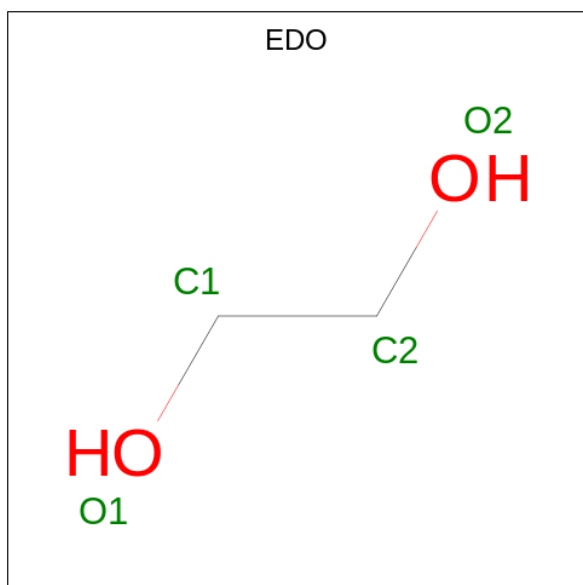
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			7	4	3		
4	A	1	Total	C	O	0	0
			7	4	3		
4	B	1	Total	C	O	0	0
			7	4	3		
4	C	1	Total	C	O	0	0
			7	4	3		

- Molecule 5 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 4 2 2	0	0
5	A	1	Total C O 4 2 2	0	0
5	A	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0

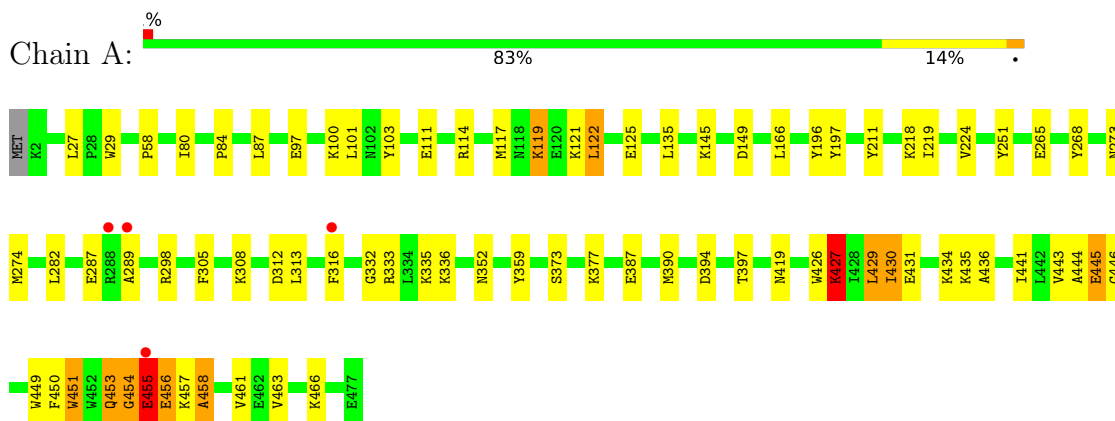
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	193	Total O 193 193	0	0
6	B	170	Total O 170 170	0	0
6	C	107	Total O 107 107	0	0
6	D	175	Total O 175 175	0	0

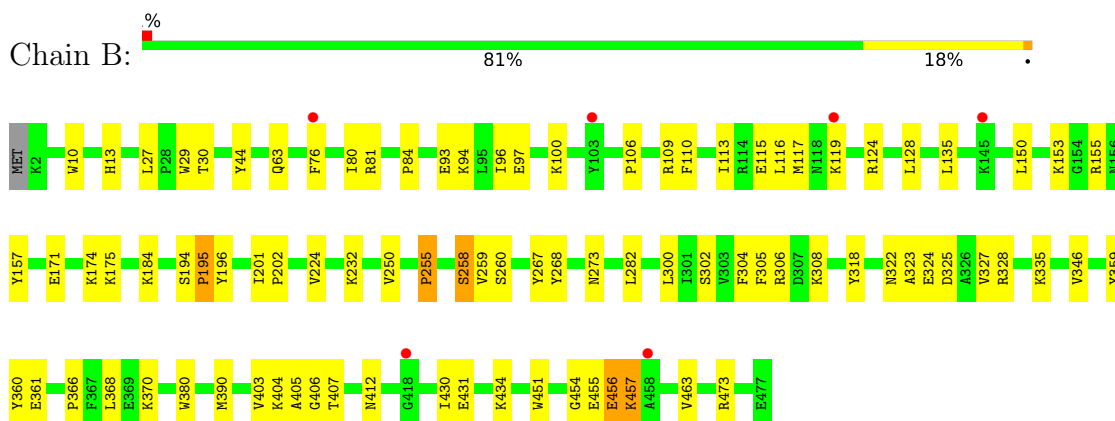
### 3 Residue-property plots

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.

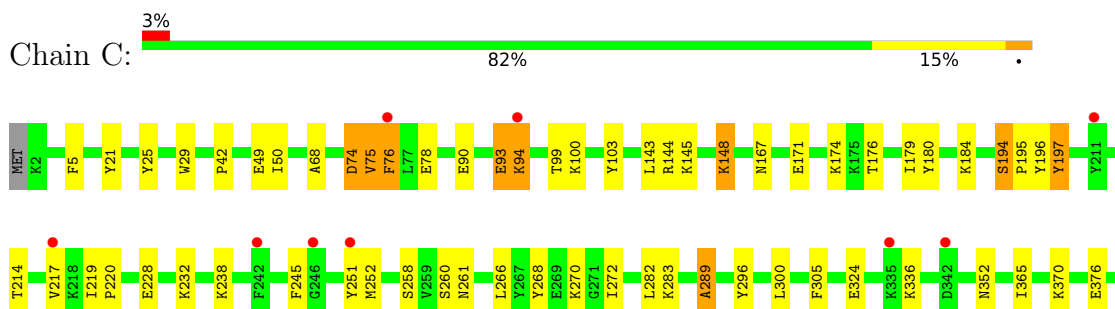
- Molecule 1: Glycoside hydrolase family 57 N-terminal domain-containing protein



- Molecule 1: Glycoside hydrolase family 57 N-terminal domain-containing protein



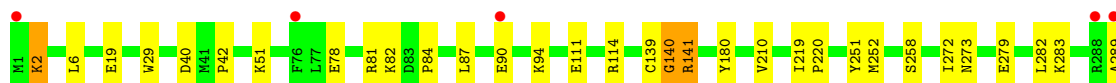
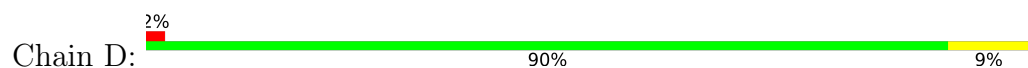
- Molecule 1: Glycoside hydrolase family 57 N-terminal domain-containing protein







- Molecule 1: Glycoside hydrolase family 57 N-terminal domain-containing protein



- Molecule 2: Cyclooctakis-(1-4)-(alpha-D-glucopyranose)



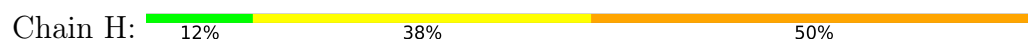
- Molecule 2: Cyclooctakis-(1-4)-(alpha-D-glucopyranose)



- Molecule 2: Cyclooctakis-(1-4)-(alpha-D-glucopyranose)



- Molecule 2: Cyclooctakis-(1-4)-(alpha-D-glucopyranose)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	40.83Å 61.53Å 195.40Å 83.94° 88.85° 89.09°	Depositor
Resolution (Å)	24.33 – 1.89 24.33 – 1.89	Depositor EDS
% Data completeness (in resolution range)	94.5 (24.33-1.89) 94.6 (24.33-1.89)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.40 (at 1.90Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, $R_{free}$	0.193 , 0.235 0.194 , 0.232	Depositor DCC
$R_{free}$ test set	7682 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.3	Xtriage
Anisotropy	0.688	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 54.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.031 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	17203	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.00% 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: PEG, GLC, EDO, GOL

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	0.80	22/4147 (0.5%)	0.67	3/5614 (0.1%)
1	B	0.66	18/4133 (0.4%)	0.60	5/5596 (0.1%)
1	C	0.71	14/4139 (0.3%)	0.60	7/5604 (0.1%)
1	D	0.54	4/4147 (0.1%)	0.58	5/5614 (0.1%)
All	All	0.68	58/16566 (0.4%)	0.61	20/22428 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	76	PHE	C-O	-8.50	1.14	1.24
1	B	456	GLU	CA-C	-7.47	1.43	1.52
1	A	458	ALA	CA-C	-7.20	1.45	1.52
1	A	455	GLU	CA-C	-7.19	1.43	1.52
1	A	427	LYS	C-O	-7.05	1.15	1.24

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	455	GLU	N-CA-C	-10.14	100.66	113.43
1	D	2	LYS	N-CA-C	-9.67	95.37	109.59
1	A	121	LYS	N-CA-C	-7.97	97.79	110.14
1	C	289	ALA	N-CA-C	7.91	122.29	108.52
1	B	455	GLU	N-CA-C	-7.46	94.84	108.65

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	4034	0	3944	53	0
1	B	4020	0	3932	69	0
1	C	4026	0	3936	63	0
1	D	4034	0	3948	41	0
2	E	88	0	65	10	0
2	F	88	0	72	4	0
2	G	88	0	72	1	0
2	H	88	0	72	6	0
3	A	18	0	23	0	0
3	B	12	0	16	1	0
3	D	6	0	8	2	0
4	A	14	0	20	0	0
4	B	7	0	10	1	0
4	C	7	0	10	3	0
5	A	12	0	18	1	0
5	B	12	0	18	12	0
5	C	4	0	6	0	0
6	A	193	0	0	12	0
6	B	170	0	0	9	0
6	C	107	0	0	7	0
6	D	175	0	0	13	0
All	All	17203	0	16170	232	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 7.

The worst 5 of 232 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:C:421:GLU:HG3	1:C:454:GLY:HA3	1.40	1.04
2:E:1:GLC:O6	2:E:2:GLC:H61	1.68	0.93
1:D:141:ARG:N	6:D:601:HOH:O	2.05	0.89
1:B:232:LYS:NZ	6:B:602:HOH:O	2.10	0.85
1:B:473:ARG:NH2	6:B:601:HOH:O	2.08	0.85

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	476/477 (100%)	466 (98%)	8 (2%)	2 (0%)	30	22
1	B	474/477 (99%)	460 (97%)	13 (3%)	1 (0%)	44	36
1	C	475/477 (100%)	463 (98%)	12 (2%)	0	100	100
1	D	476/477 (100%)	467 (98%)	8 (2%)	1 (0%)	44	36
All	All	1901/1908 (100%)	1856 (98%)	41 (2%)	4 (0%)	44	36

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	455	GLU
1	B	457	LYS
1	D	456	GLU
1	A	456	GLU

### 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	432/437 (99%)	425 (98%)	7 (2%)	58	56
1	B	430/437 (98%)	429 (100%)	1 (0%)	92	93
1	C	431/437 (99%)	424 (98%)	7 (2%)	58	56

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	432/437 (99%)	432 (100%)	0	100	100
All	All	1725/1748 (99%)	1710 (99%)	15 (1%)	75	77

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	404	LYS
1	C	396	LYS
1	C	90	GLU
1	C	455	GLU
1	C	392	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	453	GLN
1	D	69	GLN
1	D	216	ASN
1	D	156	ASN
1	C	208	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](#)

32 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	E	1	2	11,11,12	2.00	4 (36%)	15,15,17	2.62	7 (46%)
2	GLC	E	2	2	11,11,12	1.71	2 (18%)	15,15,17	2.42	8 (53%)
2	GLC	E	3	2	11,11,12	1.14	2 (18%)	15,15,17	2.75	10 (66%)
2	GLC	E	4	2	11,11,12	0.70	0	15,15,17	1.55	2 (13%)
2	GLC	E	5	2	11,11,12	0.56	0	15,15,17	1.70	3 (20%)
2	GLC	E	6	2	11,11,12	2.48	5 (45%)	15,15,17	2.29	4 (26%)
2	GLC	E	7	2	11,11,12	2.79	4 (36%)	15,15,17	2.25	3 (20%)
2	GLC	E	8	2	11,11,12	0.59	0	15,15,17	1.02	1 (6%)
2	GLC	F	1	2	11,11,12	0.74	0	15,15,17	1.06	0
2	GLC	F	2	2	11,11,12	0.54	0	15,15,17	1.48	1 (6%)
2	GLC	F	3	2	11,11,12	0.51	0	15,15,17	1.42	3 (20%)
2	GLC	F	4	2	11,11,12	0.65	0	15,15,17	1.14	3 (20%)
2	GLC	F	5	2	11,11,12	0.52	0	15,15,17	2.11	4 (26%)
2	GLC	F	6	2	11,11,12	0.76	0	15,15,17	1.25	1 (6%)
2	GLC	F	7	2	11,11,12	0.51	0	15,15,17	0.86	0
2	GLC	F	8	2	11,11,12	0.70	0	15,15,17	1.03	1 (6%)
2	GLC	G	1	2	11,11,12	0.66	0	15,15,17	0.81	1 (6%)
2	GLC	G	2	2	11,11,12	0.69	0	15,15,17	1.37	4 (26%)
2	GLC	G	3	2	11,11,12	0.71	0	15,15,17	0.90	0
2	GLC	G	4	2	11,11,12	0.58	0	15,15,17	1.42	3 (20%)
2	GLC	G	5	2	11,11,12	0.54	0	15,15,17	1.86	4 (26%)
2	GLC	G	6	2	11,11,12	0.51	0	15,15,17	1.22	2 (13%)
2	GLC	G	7	2	11,11,12	0.64	0	15,15,17	0.77	0
2	GLC	G	8	2	11,11,12	0.62	0	15,15,17	0.96	1 (6%)
2	GLC	H	1	2	11,11,12	0.55	0	15,15,17	1.54	1 (6%)
2	GLC	H	2	2	11,11,12	0.74	0	15,15,17	1.63	3 (20%)
2	GLC	H	3	2	11,11,12	0.54	0	15,15,17	1.44	3 (20%)
2	GLC	H	4	2	11,11,12	0.57	0	15,15,17	1.71	2 (13%)
2	GLC	H	5	2	11,11,12	0.50	0	15,15,17	1.87	4 (26%)
2	GLC	H	6	2	11,11,12	0.55	0	15,15,17	1.06	1 (6%)
2	GLC	H	7	2	11,11,12	0.68	0	15,15,17	0.65	0
2	GLC	H	8	2	11,11,12	0.64	0	15,15,17	1.09	2 (13%)

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	E	1	2	-	1/2/19/22	0/1/1/1
2	GLC	E	2	2	-	1/2/19/22	0/1/1/1
2	GLC	E	3	2	-	0/2/19/22	0/1/1/1
2	GLC	E	4	2	-	0/2/19/22	0/1/1/1
2	GLC	E	5	2	-	0/2/19/22	0/1/1/1
2	GLC	E	6	2	-	0/2/19/22	0/1/1/1
2	GLC	E	7	2	-	0/2/19/22	0/1/1/1
2	GLC	E	8	2	-	0/2/19/22	0/1/1/1
2	GLC	F	1	2	-	2/2/19/22	0/1/1/1
2	GLC	F	2	2	-	0/2/19/22	0/1/1/1
2	GLC	F	3	2	-	2/2/19/22	0/1/1/1
2	GLC	F	4	2	-	0/2/19/22	0/1/1/1
2	GLC	F	5	2	-	2/2/19/22	0/1/1/1
2	GLC	F	6	2	-	0/2/19/22	0/1/1/1
2	GLC	F	7	2	-	0/2/19/22	0/1/1/1
2	GLC	F	8	2	-	0/2/19/22	0/1/1/1
2	GLC	G	1	2	-	0/2/19/22	0/1/1/1
2	GLC	G	2	2	-	2/2/19/22	0/1/1/1
2	GLC	G	3	2	-	0/2/19/22	0/1/1/1
2	GLC	G	4	2	-	2/2/19/22	0/1/1/1
2	GLC	G	5	2	-	0/2/19/22	0/1/1/1
2	GLC	G	6	2	-	1/2/19/22	0/1/1/1
2	GLC	G	7	2	-	0/2/19/22	0/1/1/1
2	GLC	G	8	2	-	0/2/19/22	0/1/1/1
2	GLC	H	1	2	-	2/2/19/22	0/1/1/1
2	GLC	H	2	2	-	0/2/19/22	0/1/1/1
2	GLC	H	3	2	-	1/2/19/22	0/1/1/1
2	GLC	H	4	2	-	0/2/19/22	0/1/1/1
2	GLC	H	5	2	-	0/2/19/22	0/1/1/1
2	GLC	H	6	2	-	0/2/19/22	0/1/1/1
2	GLC	H	7	2	-	0/2/19/22	0/1/1/1
2	GLC	H	8	2	-	0/2/19/22	0/1/1/1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	7	GLC	O5-C1	-5.00	1.35	1.43
2	E	7	GLC	O3-C3	-4.82	1.31	1.43
2	E	7	GLC	O5-C5	-4.44	1.34	1.43
2	E	6	GLC	O4-C4	-4.07	1.33	1.43
2	E	6	GLC	O2-C2	-3.78	1.35	1.43

The worst 5 of 82 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	7	GLC	C1-O5-C5	6.38	120.84	112.19
2	E	6	GLC	O5-C1-C2	-6.01	101.49	110.77
2	E	3	GLC	O3-C3-C4	-5.66	97.26	110.35
2	F	5	GLC	C1-O5-C5	5.41	119.53	112.19
2	H	4	GLC	C1-O5-C5	5.07	119.06	112.19

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

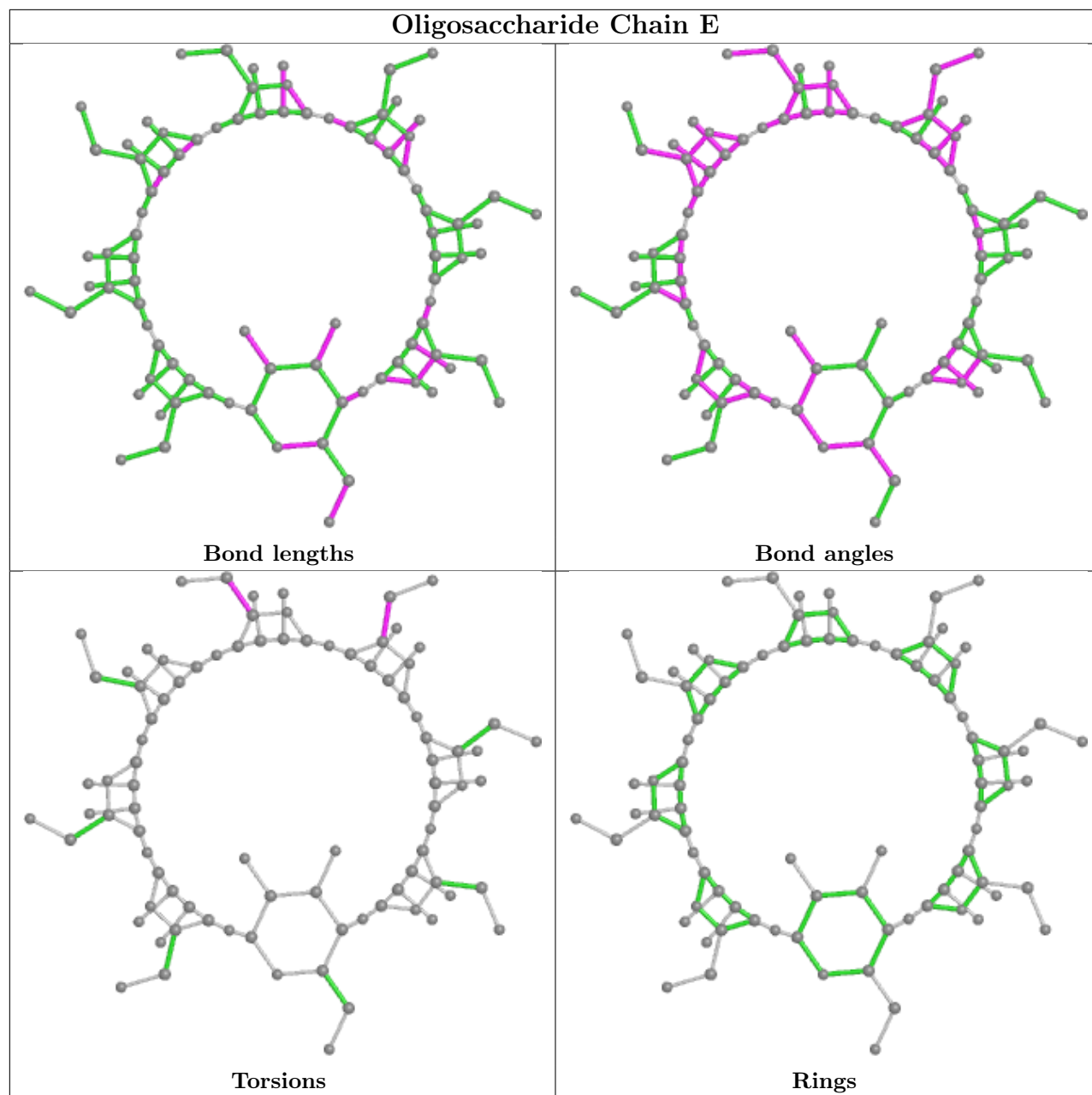
Mol	Chain	Res	Type	Atoms
2	G	2	GLC	O5-C5-C6-O6
2	G	2	GLC	C4-C5-C6-O6
2	F	1	GLC	O5-C5-C6-O6
2	H	1	GLC	O5-C5-C6-O6
2	H	1	GLC	C4-C5-C6-O6

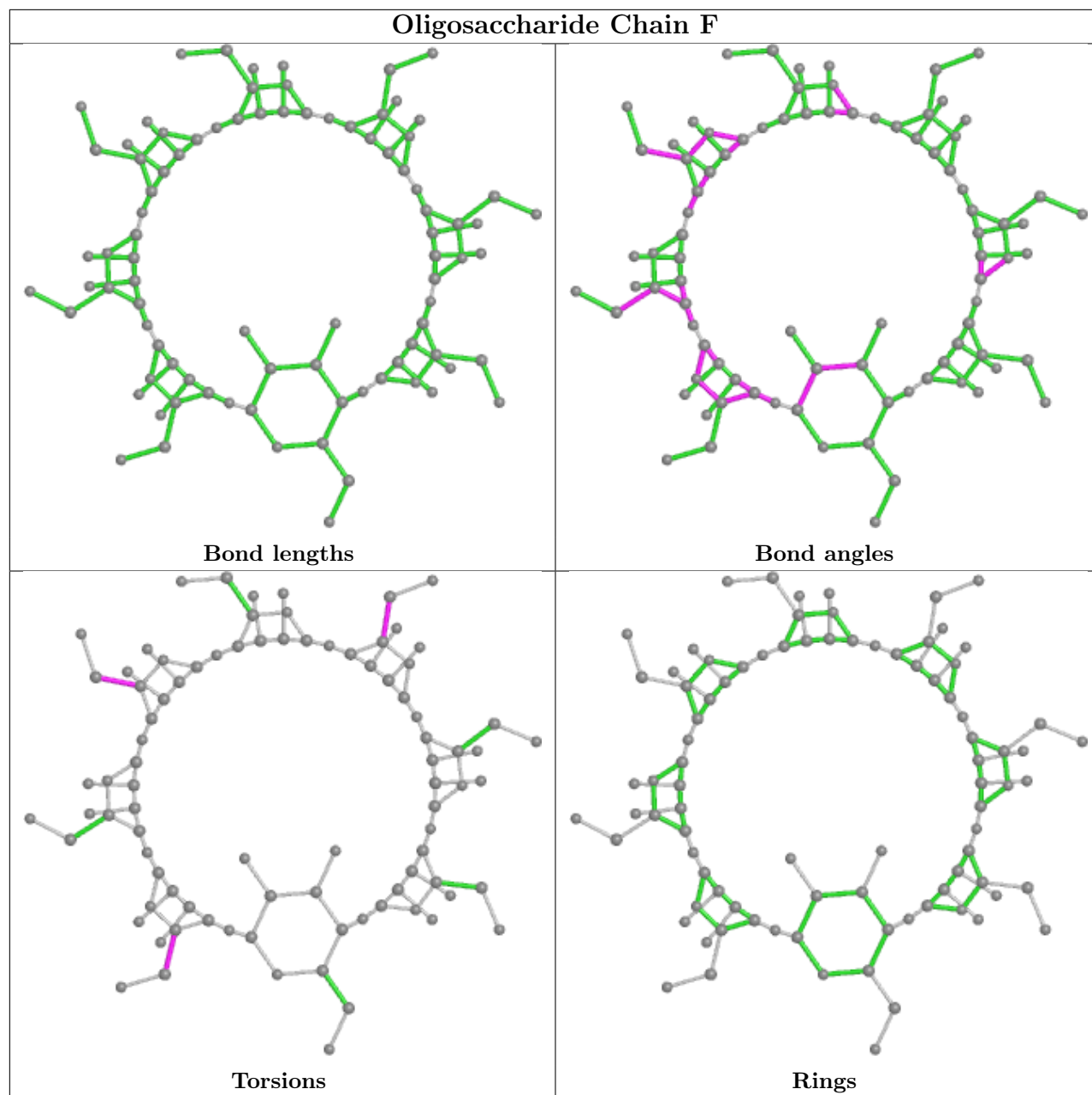
There are no ring outliers.

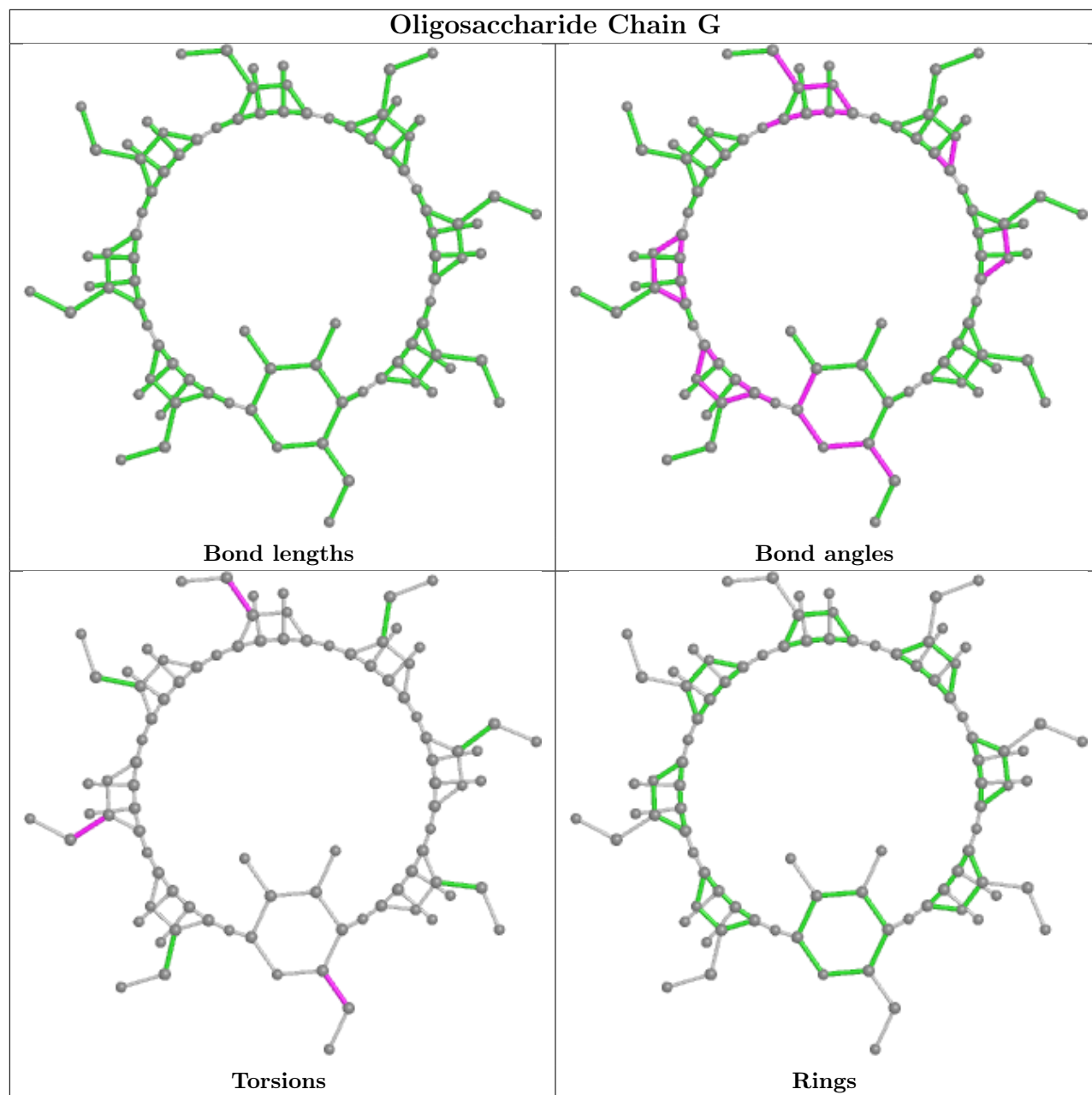
14 monomers are involved in 21 short contacts:

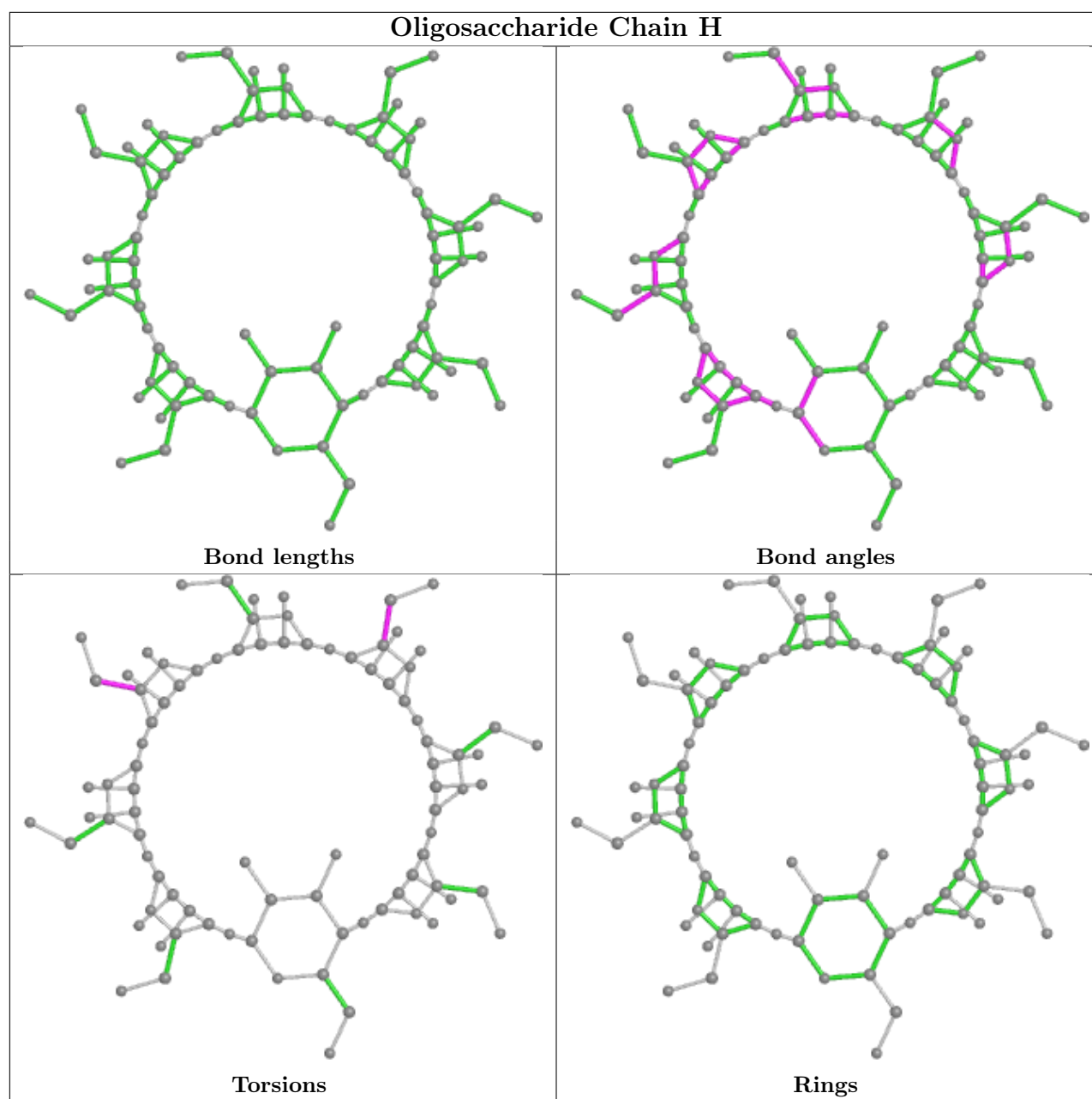
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	8	GLC	1	0
2	E	4	GLC	1	0
2	H	8	GLC	2	0
2	E	5	GLC	1	0
2	E	2	GLC	6	0
2	F	5	GLC	1	0
2	H	2	GLC	1	0
2	E	1	GLC	5	0
2	H	5	GLC	1	0
2	G	8	GLC	1	0
2	F	4	GLC	1	0
2	F	8	GLC	2	0
2	H	1	GLC	3	0
2	F	3	GLC	1	0

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](#)

17 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
5	EDO	A	506	-	3,3,3	0.55	0	2,2,2	0.29	0
3	GOL	D	501	-	5,5,5	1.06	0	5,5,5	1.34	1 (20%)
3	GOL	B	502	-	5,5,5	0.82	0	5,5,5	1.18	0
5	EDO	A	507	-	3,3,3	0.45	0	2,2,2	0.41	0
5	EDO	B	505	-	3,3,3	0.29	0	2,2,2	0.24	0
5	EDO	C	501	-	3,3,3	0.68	0	2,2,2	0.16	0
4	PEG	A	505	-	6,6,6	0.44	0	5,5,5	0.44	0
4	PEG	C	502	-	6,6,6	0.45	0	5,5,5	0.32	0
5	EDO	B	504	-	3,3,3	0.55	0	2,2,2	0.32	0
3	GOL	B	501	-	5,5,5	0.90	0	5,5,5	0.96	0
4	PEG	B	503	-	6,6,6	0.48	0	5,5,5	0.31	0
4	PEG	A	504	-	6,6,6	0.46	0	5,5,5	0.49	0
5	EDO	B	506	-	3,3,3	0.47	0	2,2,2	0.24	0
3	GOL	A	501	-	5,5,5	0.93	0	5,5,5	1.09	0
3	GOL	A	503	-	5,5,5	0.99	0	5,5,5	0.90	0
5	EDO	A	508	-	3,3,3	0.47	0	2,2,2	0.29	0
3	GOL	A	502	-	5,5,5	0.95	0	5,5,5	1.59	1 (20%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	EDO	A	506	-	-	1/1/1/1	-
3	GOL	D	501	-	-	4/4/4/4	-
3	GOL	B	502	-	-	2/4/4/4	-
5	EDO	A	507	-	-	0/1/1/1	-
5	EDO	B	505	-	-	0/1/1/1	-
5	EDO	C	501	-	-	0/1/1/1	-
4	PEG	A	505	-	-	2/4/4/4	-
4	PEG	C	502	-	-	2/4/4/4	-
5	EDO	B	504	-	-	1/1/1/1	-
3	GOL	B	501	-	-	1/4/4/4	-
4	PEG	B	503	-	-	1/4/4/4	-
4	PEG	A	504	-	-	1/4/4/4	-
5	EDO	B	506	-	-	1/1/1/1	-
3	GOL	A	501	-	-	0/4/4/4	-
3	GOL	A	503	-	-	2/4/4/4	-
5	EDO	A	508	-	-	0/1/1/1	-
3	GOL	A	502	-	-	4/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	502	GOL	C3-C2-C1	-2.77	100.93	111.70
3	D	501	GOL	C3-C2-C1	-2.23	103.03	111.70

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	502	GOL	O1-C1-C2-O2
3	B	502	GOL	C1-C2-C3-O3
3	D	501	GOL	O1-C1-C2-C3
3	D	501	GOL	C1-C2-C3-O3
3	D	501	GOL	O1-C1-C2-O2

There are no ring outliers.

7 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	501	GOL	2	0
3	B	502	GOL	1	0
5	A	507	EDO	1	0
5	B	505	EDO	8	0
4	C	502	PEG	3	0
5	B	504	EDO	4	0
4	B	503	PEG	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	476/477 (99%)	-0.15	4 (0%) 82 84	11, 29, 49, 76	2 (0%)
1	B	476/477 (99%)	0.03	6 (1%) 74 76	22, 35, 56, 74	0
1	C	476/477 (99%)	0.27	13 (2%) 56 58	17, 44, 71, 89	1 (0%)
1	D	477/477 (100%)	-0.04	8 (1%) 69 71	13, 32, 60, 84	1 (0%)
All	All	1905/1908 (99%)	0.03	31 (1%) 70 72	11, 35, 62, 89	4 (0%)

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	289	ALA	4.1
1	C	457	LYS	3.4
1	D	288	ARG	3.2
1	C	246	GLY	3.0
1	A	288	ARG	2.9

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

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

### 6.3 Carbohydrates [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	GLC	G	3	11/12	0.55	0.14	70,72,73,74	0
2	GLC	E	3	11/12	0.63	0.14	72,79,80,81	0

*Continued on next page...*

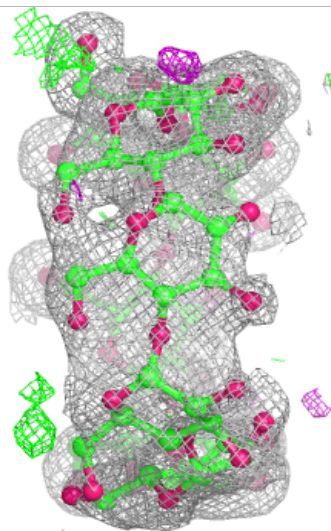
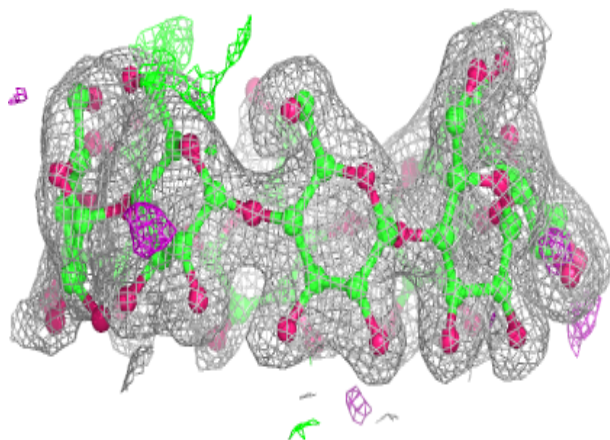
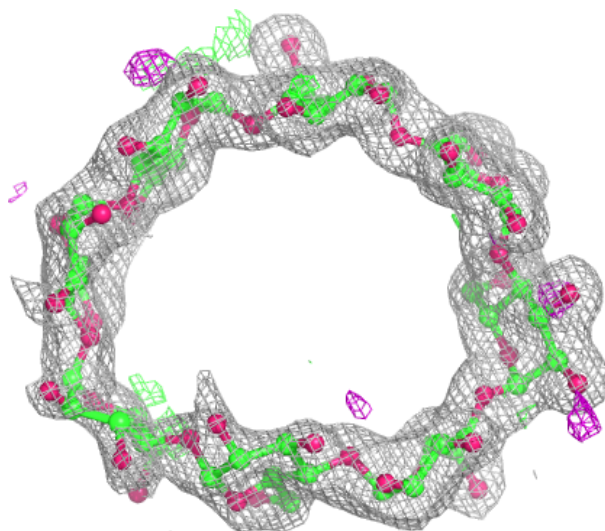
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	GLC	H	2	11/12	0.63	0.15	64,69,74,75	0
2	GLC	F	2	11/12	0.70	0.15	76,80,84,85	0
2	GLC	G	2	11/12	0.72	0.13	63,65,67,69	0
2	GLC	H	3	11/12	0.73	0.13	73,78,80,81	0
2	GLC	E	2	11/12	0.79	0.11	64,67,73,79	0
2	GLC	F	3	11/12	0.80	0.13	75,77,78,78	0
2	GLC	E	4	11/12	0.82	0.13	45,60,66,68	0
2	GLC	F	4	11/12	0.85	0.12	54,71,75,76	0
2	GLC	G	4	11/12	0.85	0.11	45,60,66,69	0
2	GLC	H	1	11/12	0.86	0.11	38,43,49,57	0
2	GLC	H	4	11/12	0.86	0.12	48,60,66,69	0
2	GLC	G	1	11/12	0.87	0.10	45,52,61,62	0
2	GLC	E	1	11/12	0.91	0.08	34,40,55,58	0
2	GLC	F	5	11/12	0.92	0.10	27,42,48,49	0
2	GLC	G	7	11/12	0.92	0.07	21,23,27,27	0
2	GLC	F	1	11/12	0.93	0.08	40,49,62,70	0
2	GLC	G	5	11/12	0.93	0.07	32,38,45,47	0
2	GLC	G	6	11/12	0.94	0.06	22,25,30,32	0
2	GLC	F	7	11/12	0.94	0.06	20,22,27,27	0
2	GLC	G	8	11/12	0.94	0.07	23,28,34,34	0
2	GLC	E	6	11/12	0.94	0.06	20,20,20,20	0
2	GLC	E	7	11/12	0.95	0.06	20,20,20,20	0
2	GLC	E	5	11/12	0.95	0.06	24,31,37,38	0
2	GLC	F	6	11/12	0.95	0.07	16,19,23,24	0
2	GLC	H	5	11/12	0.95	0.07	28,38,42,46	0
2	GLC	F	8	11/12	0.96	0.06	22,27,30,30	0
2	GLC	E	8	11/12	0.96	0.06	17,21,26,27	0
2	GLC	H	6	11/12	0.96	0.06	18,22,24,27	0
2	GLC	H	7	11/12	0.96	0.05	21,23,28,28	0
2	GLC	H	8	11/12	0.96	0.05	20,26,27,31	0

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

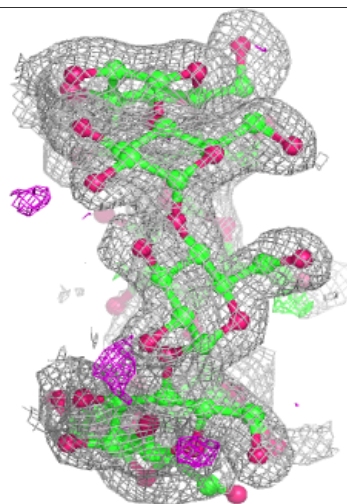
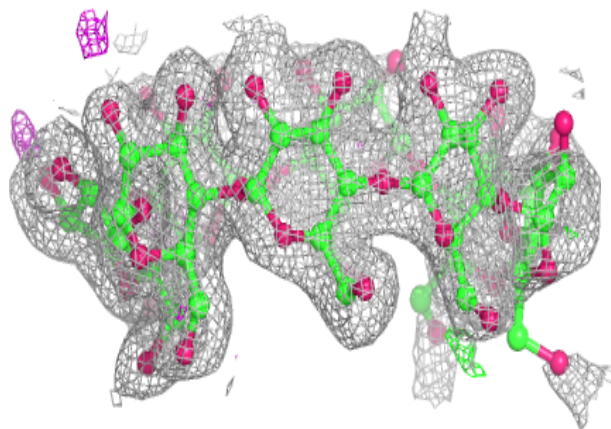
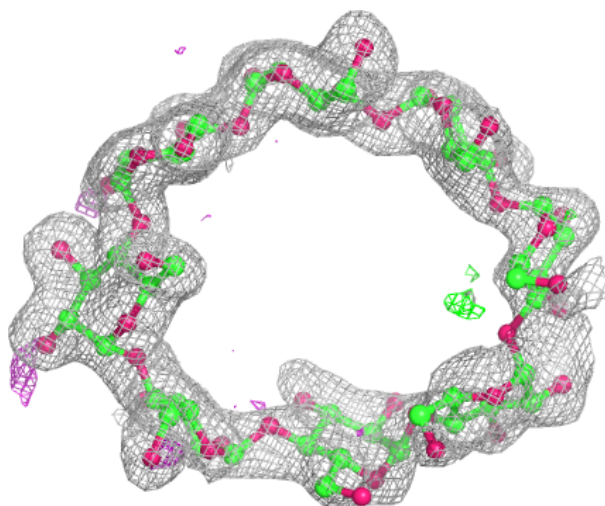
**Electron density around Chain E:**

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



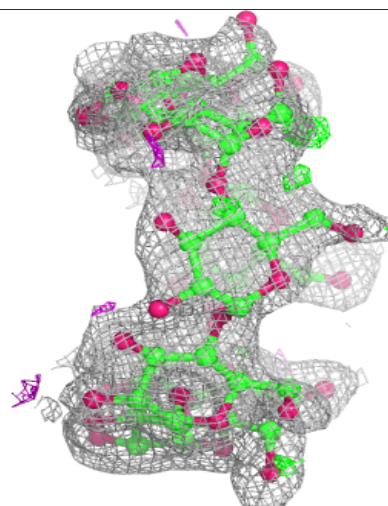
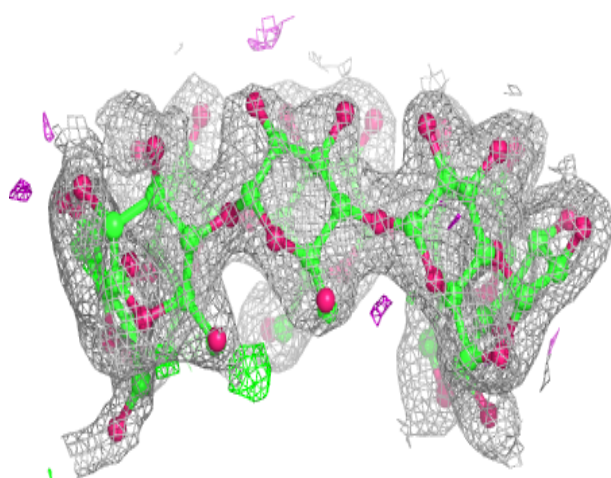
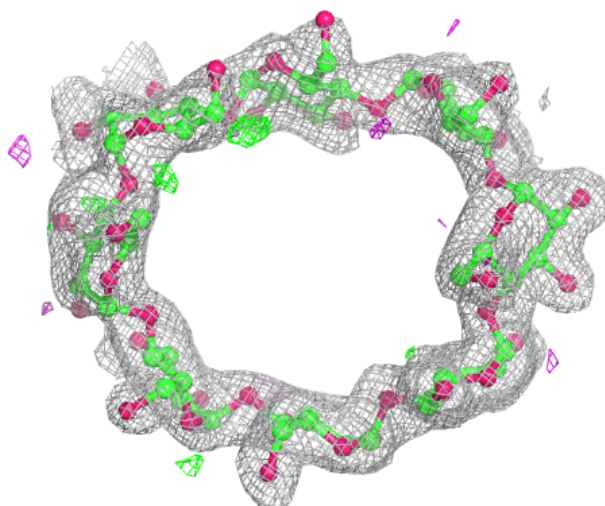
**Electron density around Chain F:**

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

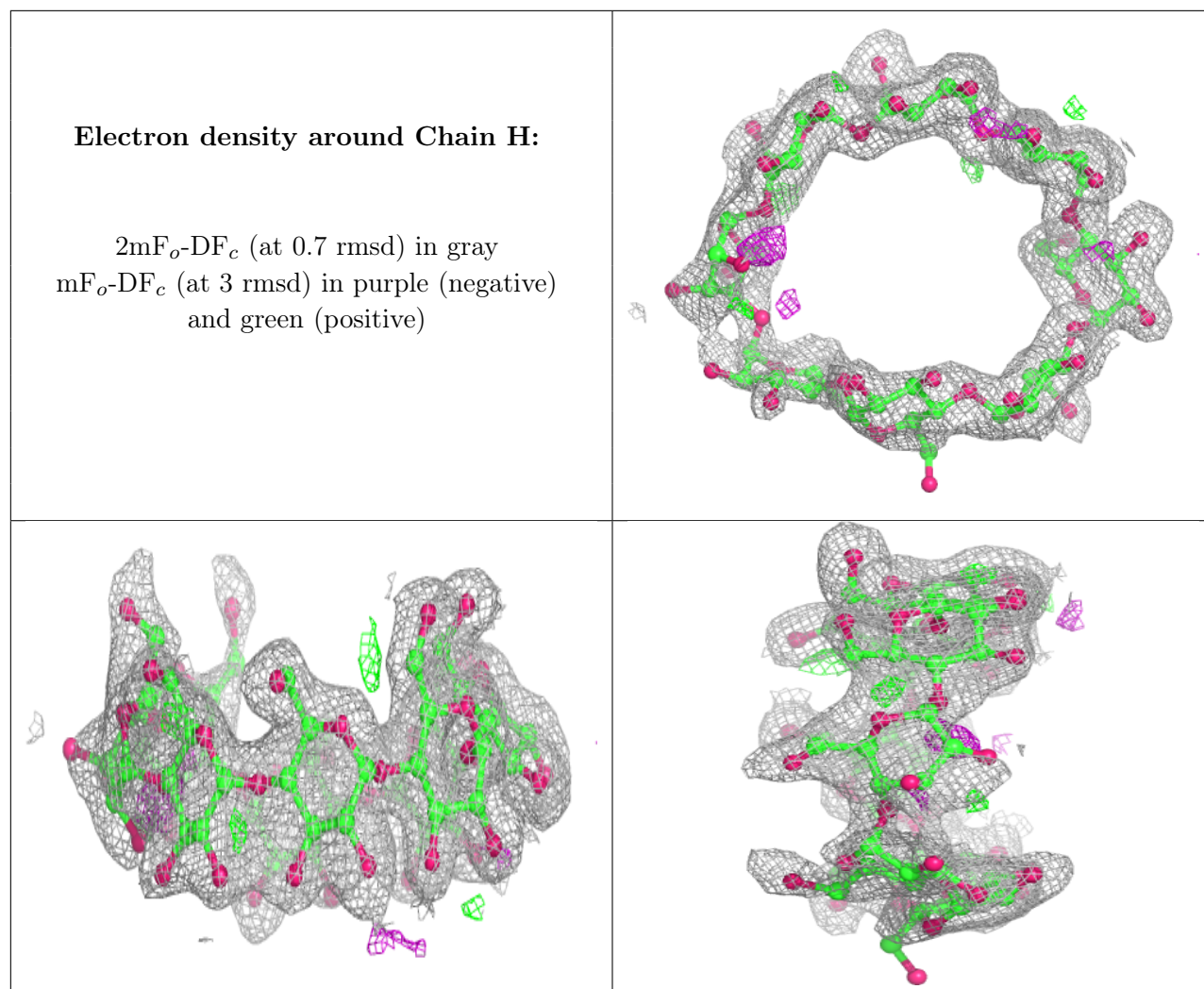


**Electron density around Chain G:**

$2mF_o-DF_c$  (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	EDO	B	504	4/4	0.60	0.26	53,58,59,60	0
5	EDO	B	506	4/4	0.67	0.25	69,70,71,73	0
4	PEG	B	503	7/7	0.68	0.19	66,67,69,70	0
4	PEG	C	502	7/7	0.73	0.15	53,55,58,59	0
4	PEG	A	504	7/7	0.79	0.15	48,53,55,57	0
5	EDO	A	506	4/4	0.79	0.15	43,49,51,51	0
5	EDO	A	507	4/4	0.80	0.15	60,64,65,65	0
5	EDO	A	508	4/4	0.80	0.17	53,53,54,55	0
4	PEG	A	505	7/7	0.81	0.15	57,64,68,69	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	GOL	B	502	6/6	0.84	0.14	39,49,52,54	0
3	GOL	B	501	6/6	0.84	0.10	53,55,58,60	0
5	EDO	B	505	4/4	0.85	0.19	36,38,39,41	0
3	GOL	A	503	6/6	0.85	0.12	39,55,57,58	0
5	EDO	C	501	4/4	0.85	0.11	29,33,35,38	0
3	GOL	A	502	6/6	0.89	0.15	27,30,33,36	0
3	GOL	A	501	6/6	0.90	0.10	47,49,53,57	0
3	GOL	D	501	6/6	0.94	0.12	26,34,39,39	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.