



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 26, 2023 – 06:53 PM EDT

PDB ID : 3G7V  
Title : Islet Amyloid Polypeptide (IAPP or Amylin) fused to Maltose Binding Protein  
Authors : Wiltzius, J.J.W.; Sawaya, M.R.; Eisenberg, D.  
Deposited on : 2009-02-10  
Resolution : 1.86 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

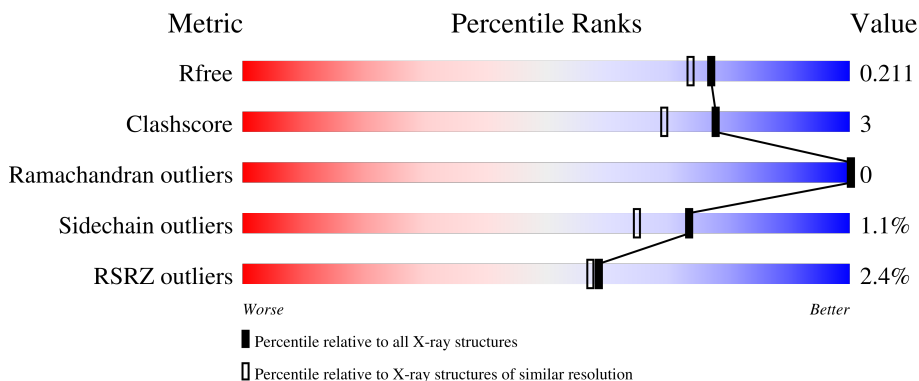
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.86 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.



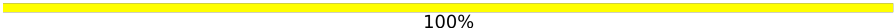
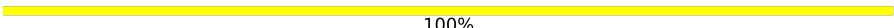

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	408	 2% 92% 6%
1	B	408	 3% 90% 7%
1	C	408	 3% 93%
1	D	408	 2% 91% 6%
2	E	2	 50% 50%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	2	 100%
2	G	2	 100%
2	H	2	 50% 50%

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
3	SO4	B	411	-	-	-	X

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 13356 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 Maltose-binding periplasmic protein, Islet amyloid polypeptide fusion protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	399	Total 3076	C 1974	N 506	O 587	S 9	0	2	0
1	B	399	Total 3079	C 1976	N 507	O 588	S 8	0	2	0
1	C	397	Total 3071	C 1972	N 505	O 586	S 8	0	3	0
1	D	397	Total 3071	C 1971	N 507	O 585	S 8	0	2	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P0AEX9
A	359	ALA	GLU	engineered mutation	UNP P0AEX9
A	362	ALA	LYS	engineered mutation	UNP P0AEX9
A	363	ALA	ASP	engineered mutation	UNP P0AEX9
A	367	ASN	-	linker	UNP P0AEX9
A	368	ALA	-	linker	UNP P0AEX9
A	369	ALA	-	linker	UNP P0AEX9
A	370	ALA	-	linker	UNP P10997
B	0	MET	-	linker	UNP P0AEX9
B	359	ALA	GLU	engineered mutation	UNP P0AEX9
B	362	ALA	LYS	engineered mutation	UNP P0AEX9
B	363	ALA	ASP	engineered mutation	UNP P0AEX9
B	367	ASN	-	linker	UNP P0AEX9
B	368	ALA	-	linker	UNP P0AEX9
B	369	ALA	-	linker	UNP P0AEX9
B	370	ALA	-	linker	UNP P10997
C	0	MET	-	linker	UNP P0AEX9
C	359	ALA	GLU	engineered mutation	UNP P0AEX9
C	362	ALA	LYS	engineered mutation	UNP P0AEX9
C	363	ALA	ASP	engineered mutation	UNP P0AEX9

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	367	ASN	-	linker	UNP P0AEX9
C	368	ALA	-	linker	UNP P0AEX9
C	369	ALA	-	linker	UNP P0AEX9
C	370	ALA	-	linker	UNP P10997
D	0	MET	-	linker	UNP P0AEX9
D	359	ALA	GLU	engineered mutation	UNP P0AEX9
D	362	ALA	LYS	engineered mutation	UNP P0AEX9
D	363	ALA	ASP	engineered mutation	UNP P0AEX9
D	367	ASN	-	linker	UNP P0AEX9
D	368	ALA	-	linker	UNP P0AEX9
D	369	ALA	-	linker	UNP P0AEX9
D	370	ALA	-	linker	UNP P10997

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	E	2	Total	C	O	0	0	0
			23	12	11			
2	F	2	Total	C	O	0	0	0
			23	12	11			
2	G	2	Total	C	O	0	0	0
			23	12	11			
2	H	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	195	Total	O	0	0
			195	195		
5	B	235	Total	O	0	0
			235	235		
5	C	226	Total	O	0	0
			226	226		

Continued on next page...

*Continued from previous page...*

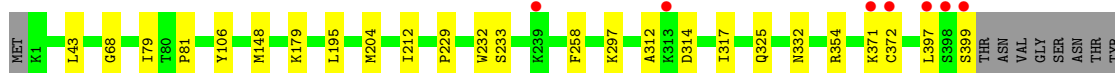
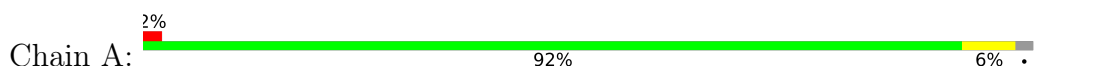
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	D	206	Total 206	O 206	0	0



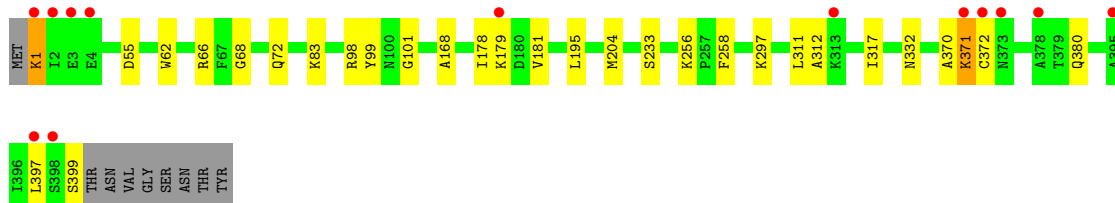
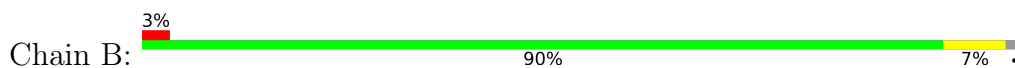
### 3 Residue-property plots [i](#)

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.

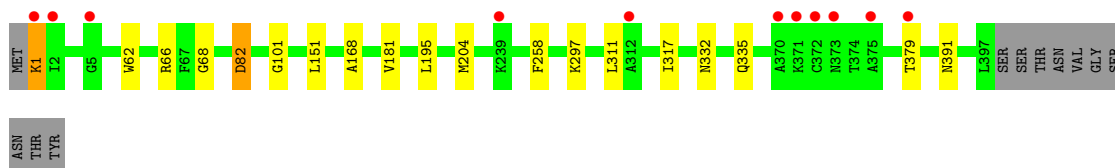
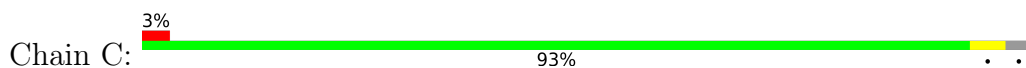
- Molecule 1: Maltose-binding periplasmic protein, Islet amyloid polypeptide fusion protein



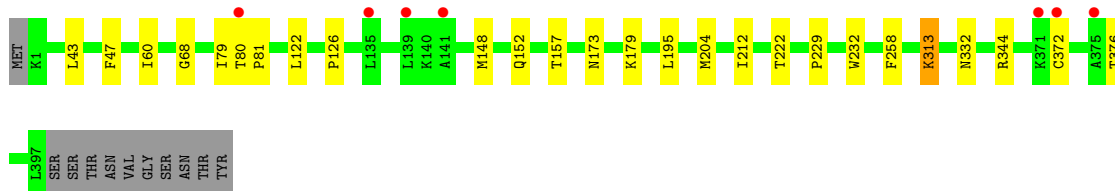
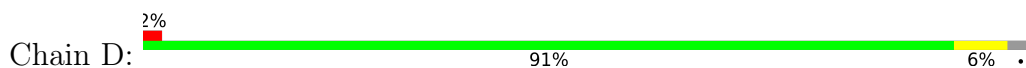
- Molecule 1: Maltose-binding periplasmic protein, Islet amyloid polypeptide fusion protein



- Molecule 1: Maltose-binding periplasmic protein, Islet amyloid polypeptide fusion protein



- Molecule 1: Maltose-binding periplasmic protein, Islet amyloid polypeptide fusion protein

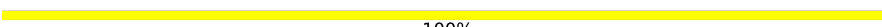


- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain E:  50% 50%

GLC1  
GLC2

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

Chain F:  100%

GLC1  
GLC2

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

Chain G:  100%

GLC1  
GLC2

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

Chain H:  50% 50%

GLC1  
GLC2

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.03Å 96.41Å 104.00Å 90.00° 90.53° 90.00°	Depositor
Resolution (Å)	73.92 – 1.86 73.89 – 1.86	Depositor EDS
% Data completeness (in resolution range)	97.9 (73.92-1.86) 97.9 (73.89-1.86)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.10 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.5.0070	Depositor
R, $R_{free}$	0.176 , 0.205 0.185 , 0.211	Depositor DCC
$R_{free}$ test set	8507 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.2	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 25.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.046 for -l,k,h 0.047 for -h,-k,l 0.460 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13356	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.15% 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: GLC, SO4, 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	0.56	0/3154	0.62	0/4284
1	B	0.55	0/3157	0.62	0/4288
1	C	0.55	0/3152	0.61	0/4282
1	D	0.55	0/3146	0.60	0/4273
All	All	0.55	0/12609	0.61	0/17127

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	A	3076	0	3051	17	0
1	B	3079	0	3054	22	0
1	C	3071	0	3046	13	0
1	D	3071	0	3044	22	0
2	E	23	0	21	0	0
2	F	23	0	21	0	0
2	G	23	0	21	0	0
2	H	23	0	21	0	0
3	A	10	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	25	0	0	0	0
3	C	25	0	0	0	0
3	D	15	0	0	0	0
4	A	6	0	8	0	0
4	B	12	0	16	0	0
4	C	6	0	8	0	0
4	D	6	0	8	0	0
5	A	195	0	0	2	0
5	B	235	0	0	4	0
5	C	226	0	0	1	0
5	D	206	0	0	5	0
All	All	13356	0	12319	73	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 3.

All (73) 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:D:148:MET:CE	1:D:212:ILE:HG22	1.90	1.00
1:D:157:THR:HG23	5:D:444:HOH:O	1.62	1.00
1:D:148:MET:HE3	1:D:212:ILE:HG22	1.57	0.86
1:D:79:ILE:HG22	1:D:81:PRO:HD3	1.58	0.85
1:D:148:MET:HE1	1:D:212:ILE:HG22	1.67	0.76
1:A:314:ASP:HB3	1:A:317:ILE:HD13	1.66	0.76
1:B:195:LEU:CD1	1:B:204:MET:HE1	2.17	0.75
1:D:376:THR:HG23	5:D:419:HOH:O	1.88	0.73
1:D:47:PHE:CG	1:D:60:ILE:HD12	2.29	0.67
1:D:148:MET:HE3	1:D:212:ILE:CG2	2.23	0.67
1:A:79:ILE:HG22	1:A:81:PRO:HD3	1.77	0.67
1:A:195:LEU:HD11	1:A:204:MET:HE1	1.77	0.66
1:A:325:GLN:NE2	1:A:397:LEU:HD22	2.11	0.65
1:C:195:LEU:HD11	1:C:204:MET:HE1	1.81	0.63
1:A:148:MET:HE1	5:A:620:HOH:O	1.99	0.61
1:D:152:GLN:O	1:D:344[B]:ARG:NH1	2.33	0.61
1:C:311:LEU:HB3	1:C:317:ILE:CD1	2.31	0.60
1:A:312:ALA:HB2	1:A:399:SER:HB3	1.85	0.59
1:C:391[A]:ASN:HD22	1:D:173:ASN:HB2	1.69	0.58
1:A:148:MET:CE	5:A:620:HOH:O	2.51	0.58
1:D:195:LEU:HD11	1:D:204:MET:HE1	1.86	0.57
1:A:148:MET:HE2	1:A:212:ILE:HG22	1.86	0.57

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:195:LEU:CD1	1:A:204:MET:HE1	2.35	0.57
1:A:148:MET:CE	1:A:212:ILE:HG22	2.35	0.56
1:A:68:GLY:HA3	1:A:332:ASN:O	2.06	0.56
1:C:151:LEU:HD11	1:C:204:MET:HE3	1.91	0.53
1:B:312:ALA:HB2	1:B:399:SER:HB3	1.91	0.53
1:B:68:GLY:HA3	1:B:332:ASN:O	2.09	0.52
1:A:179:LYS:HE2	1:A:372:CYS:HA	1.91	0.52
1:C:101:GLY:HA3	1:C:379:THR:OG1	2.09	0.52
1:D:179:LYS:HE2	1:D:372:CYS:HA	1.92	0.52
1:B:72[A]:GLN:NE2	1:B:99:TYR:OH	2.43	0.52
1:B:195:LEU:HD12	1:B:204:MET:HE1	1.92	0.51
1:B:195:LEU:HD11	1:B:204:MET:HE1	1.91	0.50
1:B:397:LEU:C	1:B:397:LEU:HD13	2.33	0.49
1:B:311:LEU:HB3	1:B:317:ILE:CD1	2.42	0.49
1:D:68:GLY:HA3	1:D:332:ASN:O	2.13	0.49
1:B:195:LEU:CD1	1:B:204:MET:CE	2.91	0.48
1:B:62:TRP:CD1	1:B:66:ARG:HG3	2.49	0.48
1:A:43:LEU:C	1:A:43:LEU:HD12	2.35	0.47
1:B:371:LYS:HD2	1:B:371:LYS:H	1.79	0.47
1:A:233:SER:OG	1:A:297:LYS:HD3	2.15	0.47
1:B:178:ILE:HG13	1:B:370:ALA:HB3	1.97	0.47
1:B:256:LYS:HE2	5:B:681:HOH:O	2.15	0.46
1:A:79:ILE:HD12	1:A:106:TYR:CE2	2.50	0.46
1:B:1:LYS:NZ	1:B:1:LYS:HB3	2.30	0.46
1:C:62:TRP:CD1	1:C:66:ARG:HG3	2.50	0.46
1:B:98:ARG:HB3	5:B:911:HOH:O	2.15	0.45
1:C:311:LEU:HB3	1:C:317:ILE:HD13	1.98	0.45
1:D:195:LEU:CD1	1:D:204:MET:HE1	2.47	0.45
1:A:229:PRO:HA	1:A:232:TRP:CE2	2.51	0.45
1:D:43:LEU:HD12	1:D:43:LEU:C	2.37	0.45
1:D:229:PRO:HA	1:D:232:TRP:CE2	2.52	0.45
1:C:1:LYS:NZ	1:C:1:LYS:HB3	2.32	0.43
1:B:195:LEU:HD12	1:B:204:MET:CE	2.48	0.43
1:C:68:GLY:HA3	1:C:332:ASN:O	2.18	0.43
1:C:195:LEU:CD1	1:C:204:MET:HE1	2.46	0.43
1:D:376:THR:HG21	5:D:774:HOH:O	2.16	0.43
1:A:325:GLN:HE21	1:A:397:LEU:HD22	1.79	0.43
1:B:168:ALA:O	1:B:181:VAL:HA	2.19	0.43
1:B:380:GLN:HA	5:B:911:HOH:O	2.18	0.43
1:B:179:LYS:CG	1:B:372:CYS:HA	2.48	0.43
1:C:335[A]:GLN:NE2	1:C:335[A]:GLN:H	2.17	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:313:LYS:NZ	5:D:445:HOH:O	2.52	0.42
1:D:148:MET:CE	1:D:212:ILE:CG2	2.77	0.42
1:D:148:MET:HB2	1:D:222:THR:HG21	2.01	0.42
1:B:179:LYS:CE	1:B:372:CYS:HA	2.50	0.41
1:C:82:ASP:HB2	5:C:697:HOH:O	2.21	0.41
1:D:80:THR:HG22	5:D:745:HOH:O	2.20	0.41
1:C:168:ALA:O	1:C:181:VAL:HA	2.21	0.41
1:B:101:GLY:HA2	5:B:911:HOH:O	2.21	0.40
1:D:122:LEU:HD21	1:D:126:PRO:HD3	2.04	0.40
1:B:233:SER:OG	1:B:297:LYS:HD3	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

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

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	399/408 (98%)	395 (99%)	4 (1%)	0	100	100
1	B	399/408 (98%)	396 (99%)	3 (1%)	0	100	100
1	C	398/408 (98%)	394 (99%)	4 (1%)	0	100	100
1	D	397/408 (97%)	393 (99%)	4 (1%)	0	100	100
All	All	1593/1632 (98%)	1578 (99%)	15 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	317/323 (98%)	314 (99%)	3 (1%)	78	72
1	B	317/323 (98%)	312 (98%)	5 (2%)	62	49
1	C	316/323 (98%)	312 (99%)	4 (1%)	69	58
1	D	315/323 (98%)	313 (99%)	2 (1%)	86	83
All	All	1265/1292 (98%)	1251 (99%)	14 (1%)	73	65

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	258	PHE
1	A	354	ARG
1	A	371	LYS
1	B	1	LYS
1	B	55	ASP
1	B	83	LYS
1	B	258	PHE
1	B	371	LYS
1	C	1	LYS
1	C	82	ASP
1	C	258	PHE
1	C	297	LYS
1	D	258	PHE
1	D	313	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	325	GLN
1	A	391	ASN
1	B	391	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](#)

8 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GLC	E	1	2	12,12,12	0.75	0	17,17,17	1.91	5 (29%)
2	GLC	E	2	2	11,11,12	0.60	0	15,15,17	0.85	0
2	GLC	F	1	2	12,12,12	0.75	0	17,17,17	1.64	4 (23%)
2	GLC	F	2	2	11,11,12	0.54	0	15,15,17	1.05	1 (6%)
2	GLC	G	1	2	12,12,12	0.88	0	17,17,17	1.93	4 (23%)
2	GLC	G	2	2	11,11,12	0.61	0	15,15,17	1.08	1 (6%)
2	GLC	H	1	2	12,12,12	0.69	0	17,17,17	1.74	4 (23%)
2	GLC	H	2	2	11,11,12	0.58	0	15,15,17	0.95	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	E	1	2	-	0/2/22/22	0/1/1/1
2	GLC	E	2	2	-	0/2/19/22	0/1/1/1
2	GLC	F	1	2	-	0/2/22/22	0/1/1/1
2	GLC	F	2	2	-	0/2/19/22	0/1/1/1
2	GLC	G	1	2	-	0/2/22/22	0/1/1/1
2	GLC	G	2	2	-	0/2/19/22	0/1/1/1
2	GLC	H	1	2	-	0/2/22/22	0/1/1/1
2	GLC	H	2	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	1	GLC	O5-C1-C2	5.10	119.38	110.28
2	G	1	GLC	O5-C1-C2	4.55	118.40	110.28
2	H	1	GLC	O1-C1-O5	4.34	123.39	110.38
2	G	1	GLC	C1-C2-C3	-3.65	102.75	110.31
2	F	1	GLC	O1-C1-O5	3.61	121.21	110.38
2	F	1	GLC	O5-C1-C2	3.42	116.39	110.28
2	H	1	GLC	O5-C1-C2	3.41	116.37	110.28
2	G	1	GLC	O1-C1-O5	3.38	120.51	110.38
2	G	2	GLC	C1-O5-C5	3.29	116.66	112.19
2	E	1	GLC	O1-C1-O5	2.98	119.31	110.38
2	F	2	GLC	C1-O5-C5	2.86	116.07	112.19
2	E	1	GLC	C1-O5-C5	-2.83	108.32	113.66
2	F	1	GLC	C1-C2-C3	-2.82	104.47	110.31
2	E	1	GLC	C1-C2-C3	-2.77	104.58	110.31
2	G	1	GLC	C1-O5-C5	-2.63	108.71	113.66
2	H	1	GLC	C4-C3-C2	2.35	114.93	110.82
2	E	1	GLC	O1-C1-C2	2.35	115.64	109.03
2	F	1	GLC	C1-O5-C5	-2.27	109.38	113.66
2	H	1	GLC	O1-C1-C2	2.24	115.35	109.03

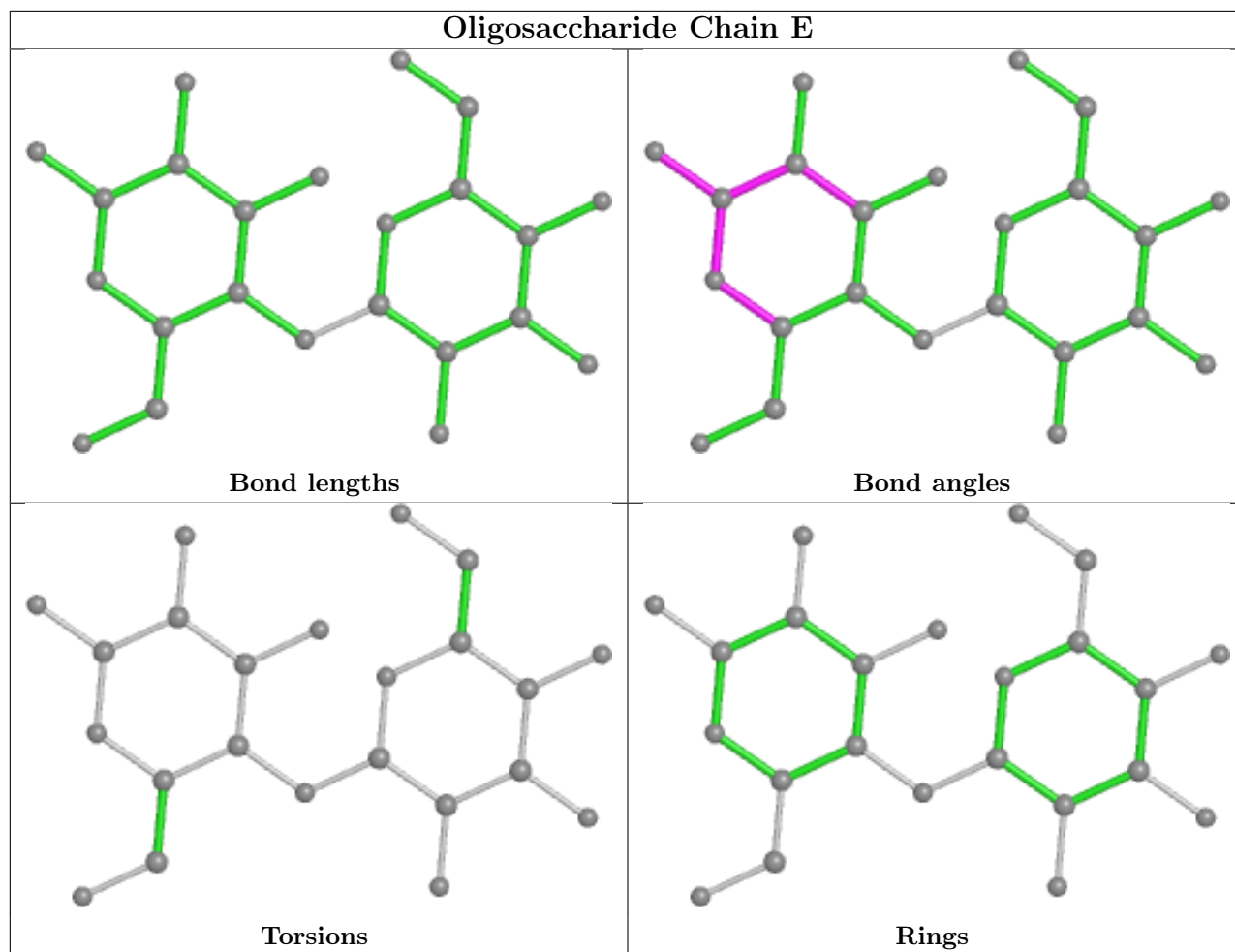
There are no chirality outliers.

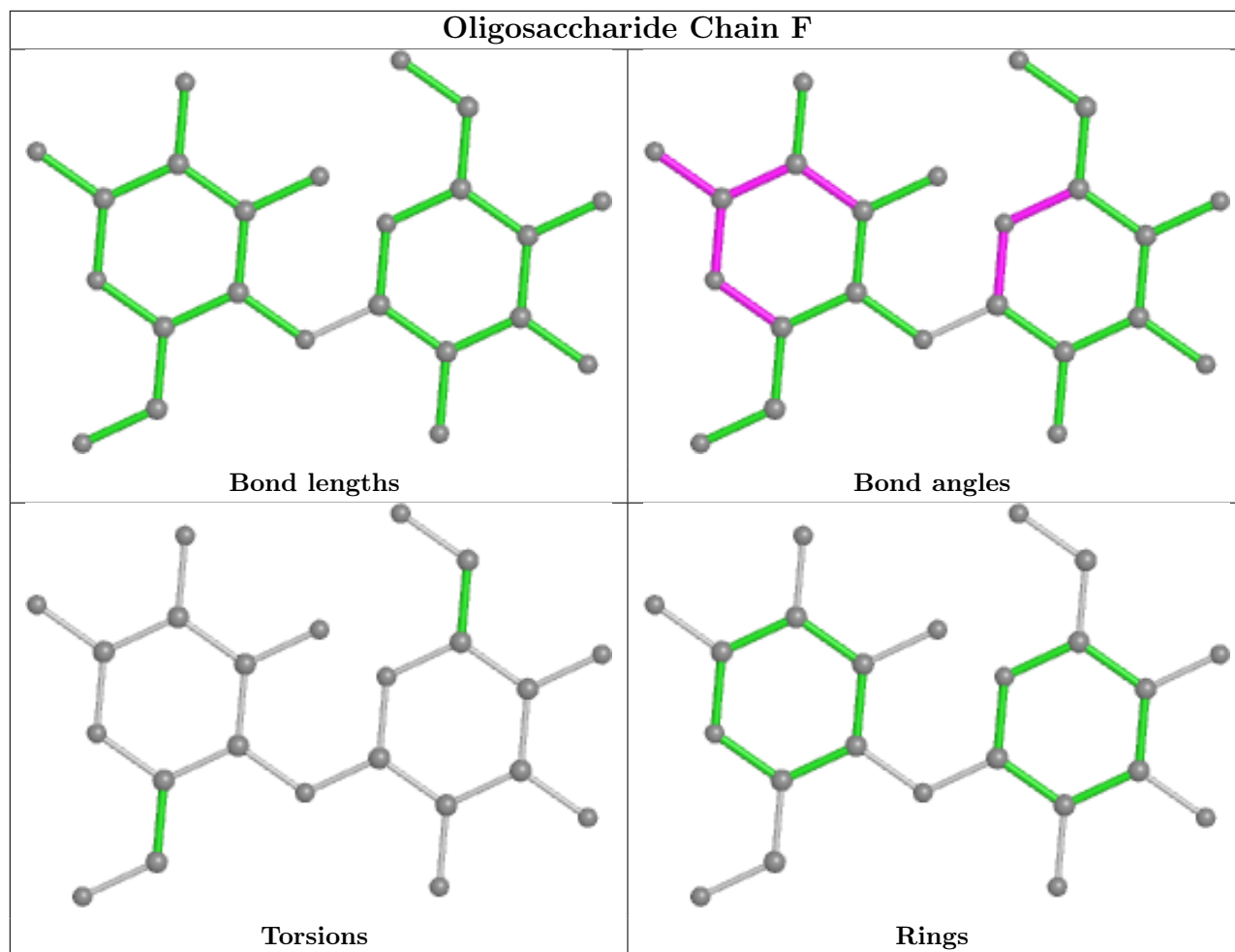
There are no torsion outliers.

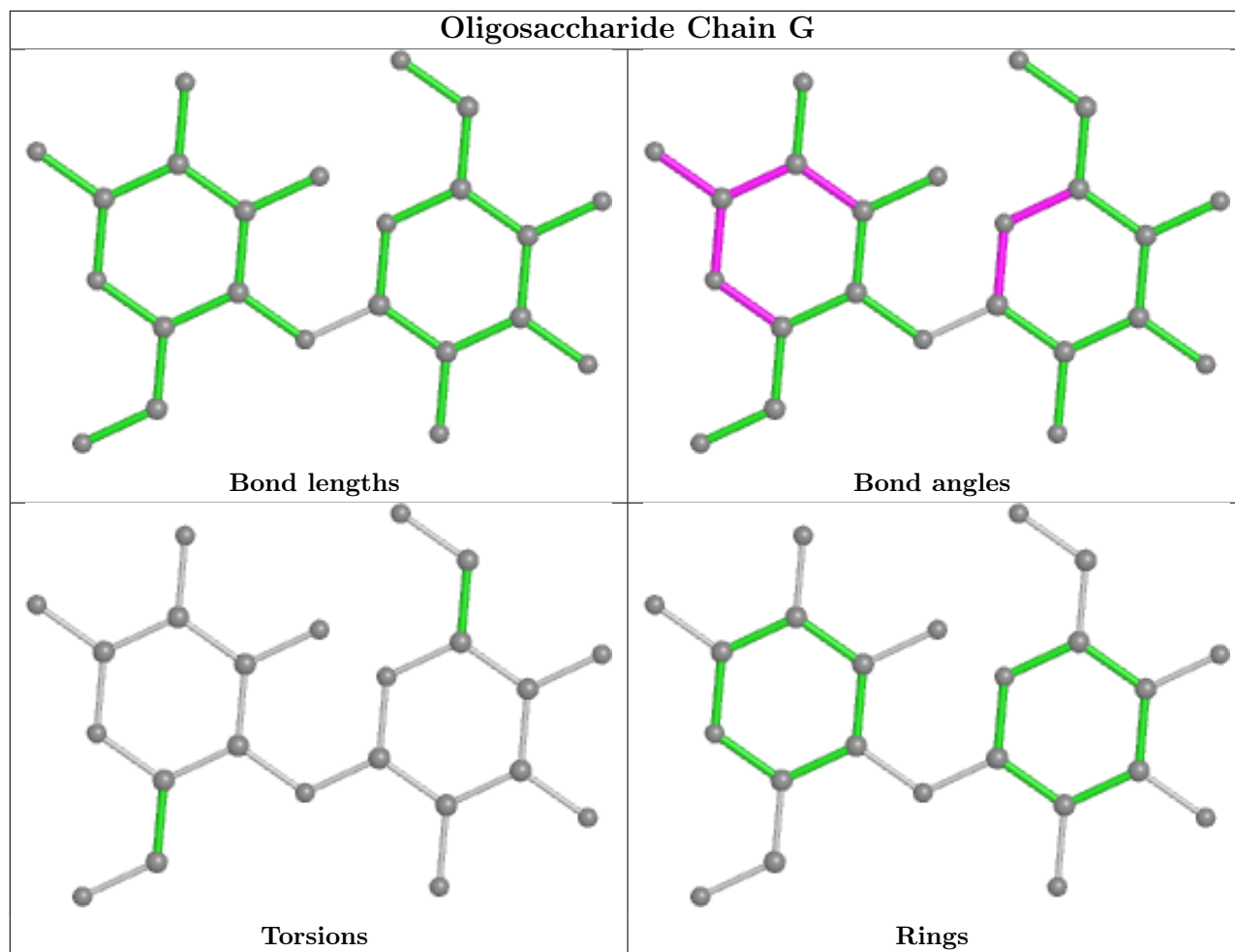
There are no ring outliers.

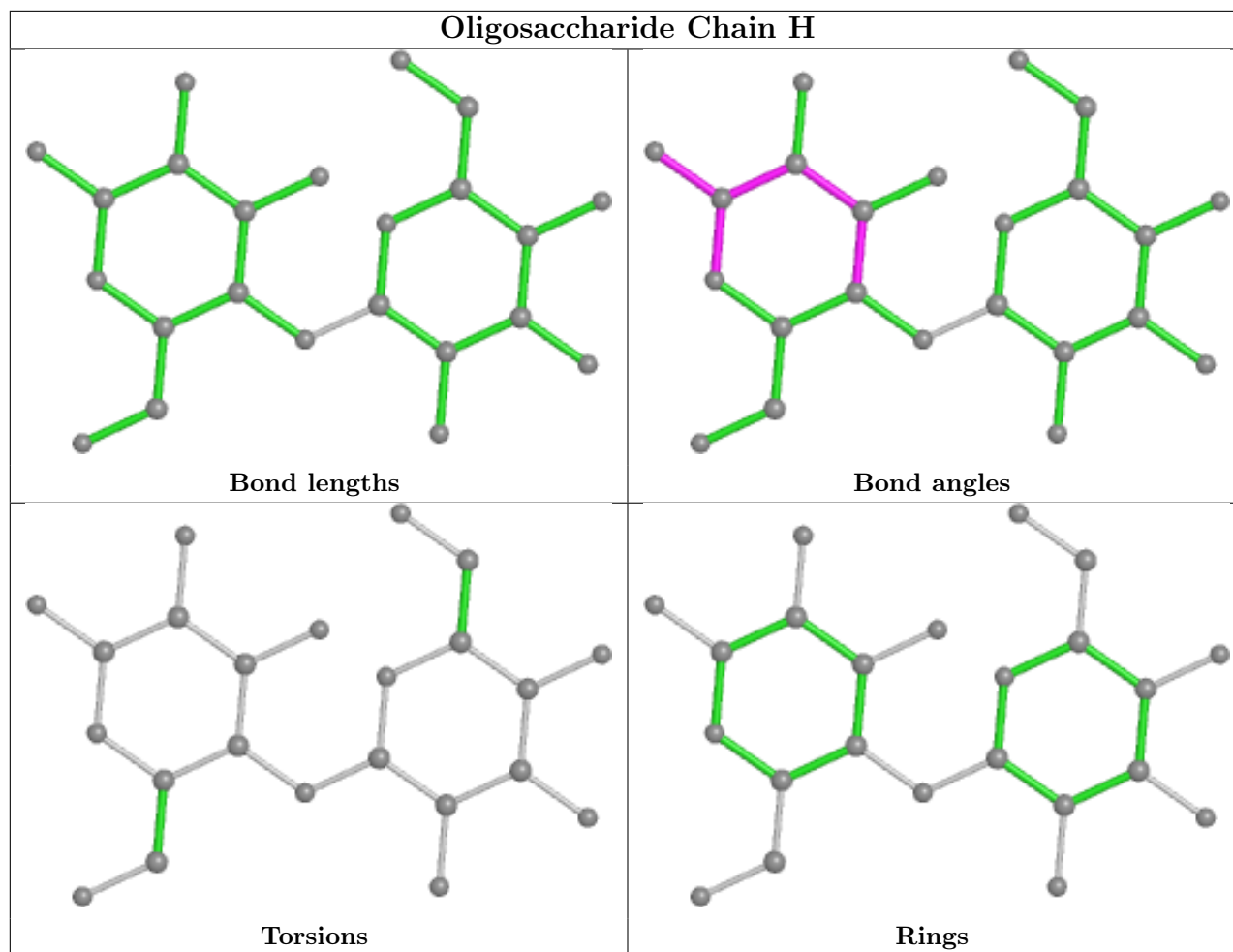
No monomer is involved in short contacts.

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









## 5.6 Ligand geometry [i](#)

20 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	$\# Z  > 2$
3	SO4	D	409	-	4,4,4	0.17	0	6,6,6	0.36	0
4	GOL	D	411	-	5,5,5	0.36	0	5,5,5	0.27	0
3	SO4	B	410	-	4,4,4	0.18	0	6,6,6	0.17	0
4	GOL	B	414	-	5,5,5	0.36	0	5,5,5	0.15	0
4	GOL	B	413	-	5,5,5	0.37	0	5,5,5	0.19	0
3	SO4	D	410	-	4,4,4	0.15	0	6,6,6	0.15	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	SO4	C	409	-	4,4,4	0.25	0	6,6,6	0.31	0
3	SO4	A	408	-	4,4,4	0.30	0	6,6,6	0.44	0
3	SO4	B	408	-	4,4,4	0.22	0	6,6,6	0.23	0
3	SO4	B	409	-	4,4,4	0.40	0	6,6,6	0.30	0
3	SO4	D	408	-	4,4,4	0.17	0	6,6,6	0.30	0
3	SO4	C	412	-	4,4,4	0.13	0	6,6,6	0.14	0
3	SO4	B	411	-	4,4,4	0.17	0	6,6,6	0.12	0
3	SO4	B	412	-	4,4,4	0.13	0	6,6,6	0.16	0
3	SO4	C	411	-	4,4,4	0.20	0	6,6,6	0.24	0
3	SO4	C	408	-	4,4,4	0.18	0	6,6,6	0.18	0
3	SO4	C	410	-	4,4,4	0.16	0	6,6,6	0.21	0
4	GOL	C	413	-	5,5,5	0.40	0	5,5,5	0.63	0
3	SO4	A	409	-	4,4,4	0.15	0	6,6,6	0.11	0
4	GOL	A	410	-	5,5,5	0.39	0	5,5,5	0.25	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	C	413	-	-	2/4/4/4	-
4	GOL	D	411	-	-	2/4/4/4	-
4	GOL	B	414	-	-	0/4/4/4	-
4	GOL	A	410	-	-	2/4/4/4	-
4	GOL	B	413	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	410	GOL	O1-C1-C2-C3
4	B	413	GOL	O1-C1-C2-C3
4	B	413	GOL	O1-C1-C2-O2
4	C	413	GOL	C1-C2-C3-O3
4	D	411	GOL	O1-C1-C2-C3
4	A	410	GOL	O1-C1-C2-O2
4	C	413	GOL	O2-C2-C3-O3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	D	411	GOL	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

## 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

### 6.1 Protein, DNA and RNA chains

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	399/408 (97%)	0.20	7 (1%) 68 68	9, 18, 31, 68	0
1	B	399/408 (97%)	0.26	13 (3%) 46 44	11, 18, 33, 69	0
1	C	397/408 (97%)	0.19	11 (2%) 53 52	10, 17, 32, 65	0
1	D	397/408 (97%)	0.23	7 (1%) 68 68	10, 17, 30, 52	0
All	All	1592/1632 (97%)	0.22	38 (2%) 59 57	9, 18, 32, 69	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	372	CYS	9.4
1	D	372	CYS	8.7
1	C	372	CYS	7.4
1	B	1	LYS	6.1
1	C	371	LYS	5.9
1	B	3	GLU	5.9
1	C	2	ILE	5.8
1	C	1	LYS	5.6
1	B	2	ILE	5.5
1	A	399	SER	4.9
1	C	5	GLY	4.3
1	C	375	ALA	4.3
1	D	135	LEU	4.0
1	D	141	ALA	3.8
1	D	371	LYS	3.8
1	D	139	LEU	3.6
1	C	370	ALA	3.5
1	B	397	LEU	3.5
1	A	372	CYS	3.4
1	A	313	LYS	3.4
1	A	371	LYS	3.3

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	398	SER	3.1
1	C	312	ALA	3.0
1	D	375	ALA	2.9
1	B	179	LYS	2.7
1	A	397	LEU	2.7
1	C	373	ASN	2.7
1	C	379	THR	2.6
1	B	371	LYS	2.6
1	B	398	SER	2.5
1	D	80	THR	2.4
1	A	239	LYS	2.4
1	B	373	ASN	2.4
1	C	239	LYS	2.2
1	B	395	ALA	2.2
1	B	4	GLU	2.1
1	B	313	LYS	2.1
1	B	378	ALA	2.0

## 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](#)

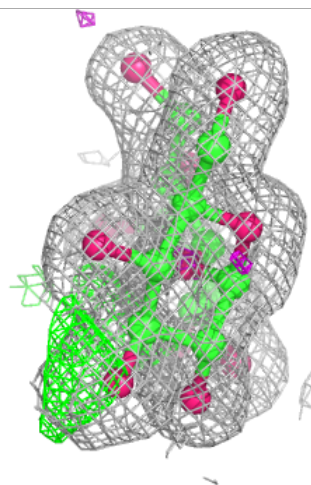
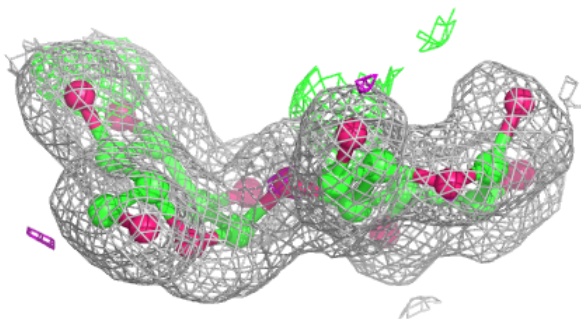
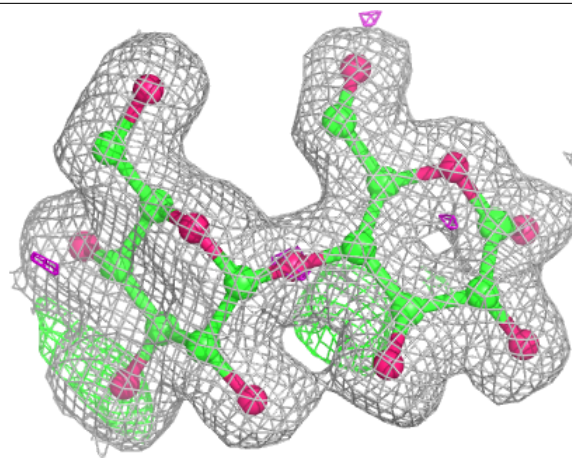
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	H	1	12/12	0.89	0.11	8,18,23,28	0
2	GLC	G	1	12/12	0.92	0.10	10,18,23,24	0
2	GLC	E	1	12/12	0.93	0.09	8,17,24,31	0
2	GLC	E	2	11/12	0.94	0.11	15,16,17,18	0
2	GLC	G	2	11/12	0.94	0.11	13,16,18,18	0
2	GLC	F	1	12/12	0.94	0.09	11,20,23,24	0
2	GLC	F	2	11/12	0.95	0.12	15,17,19,19	0
2	GLC	H	2	11/12	0.96	0.09	14,16,18,18	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.

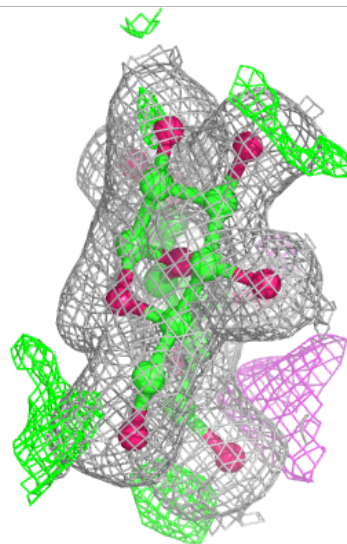
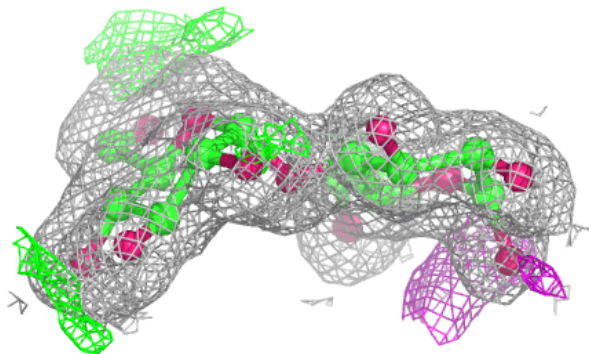
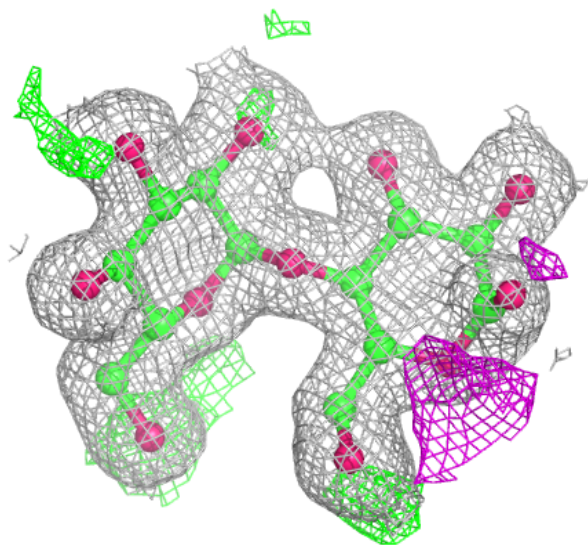
**Electron density around Chain E:**

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



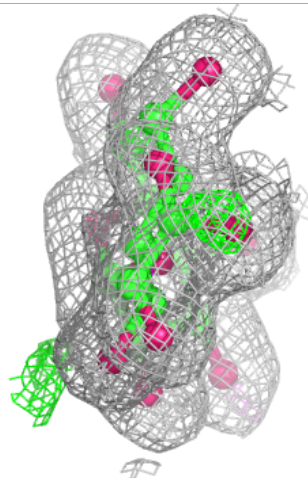
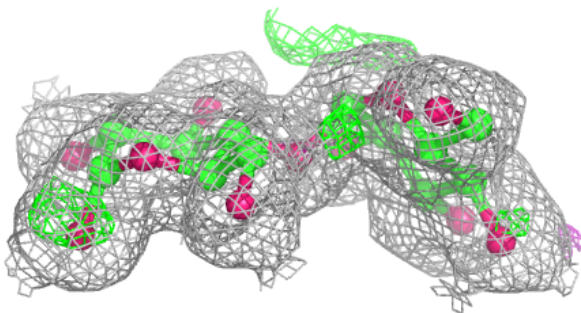
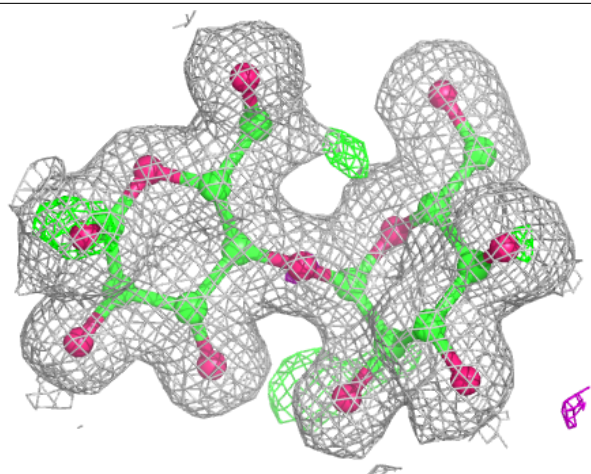
**Electron density around Chain F:**

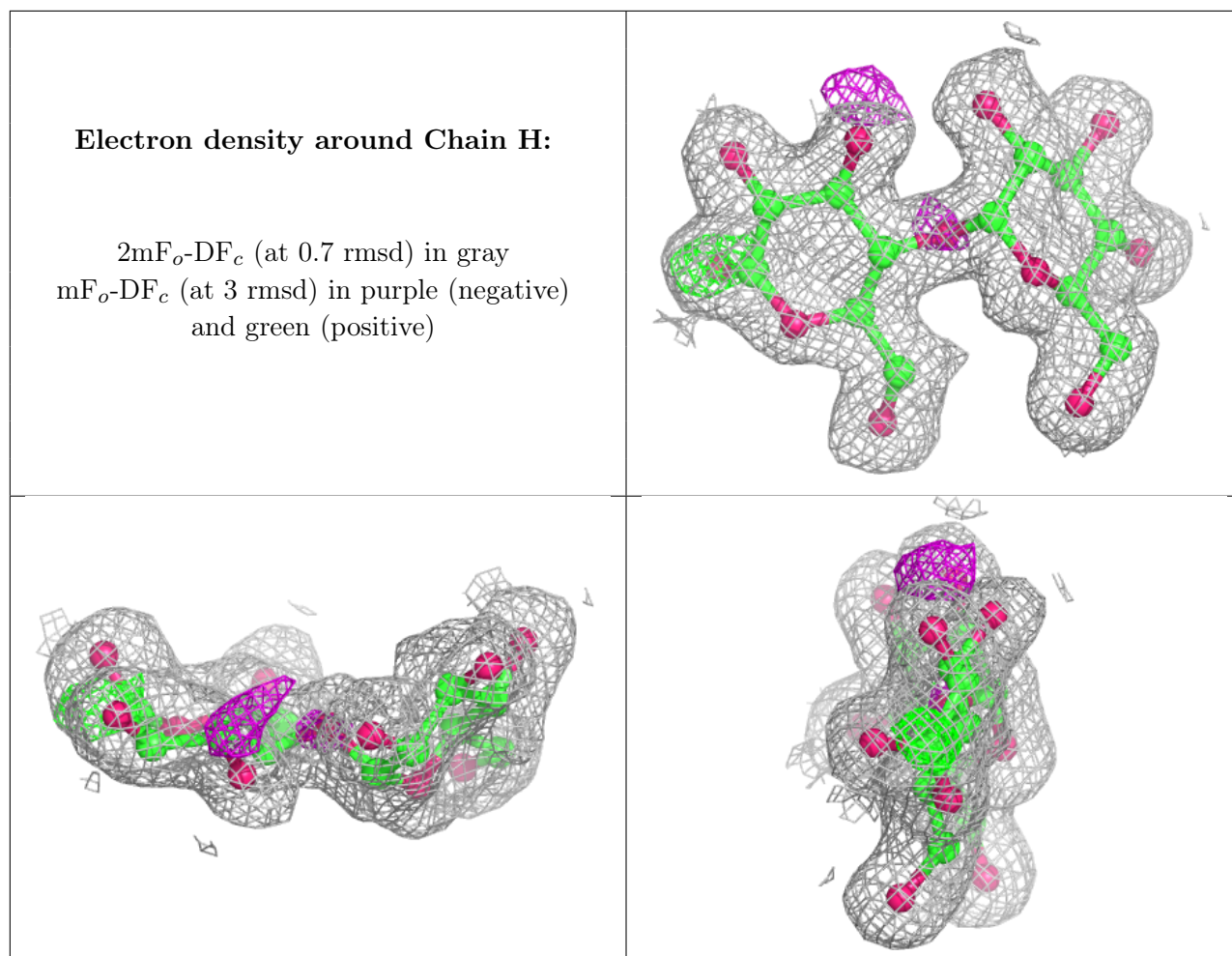
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**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
4	GOL	B	413	6/6	0.59	0.33	66,67,67,68	0
4	GOL	D	411	6/6	0.67	0.30	60,62,62,63	0
4	GOL	C	413	6/6	0.70	0.18	55,55,56,56	0
3	SO4	B	411	5/5	0.71	0.40	85,86,86,87	0
4	GOL	A	410	6/6	0.76	0.21	56,57,59,59	0
3	SO4	B	409	5/5	0.77	0.26	35,38,42,46	0
3	SO4	A	409	5/5	0.77	0.27	67,69,69,70	0
3	SO4	C	410	5/5	0.84	0.23	67,67,69,69	0
3	SO4	D	409	5/5	0.84	0.14	57,60,63,65	0
3	SO4	C	412	5/5	0.85	0.16	77,79,79,80	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	SO4	B	408	5/5	0.85	0.23	52,56,59,59	0
3	SO4	C	411	5/5	0.85	0.22	49,56,57,58	0
3	SO4	C	408	5/5	0.86	0.22	64,65,65,68	0
3	SO4	B	410	5/5	0.86	0.21	66,67,68,70	0
3	SO4	B	412	5/5	0.86	0.17	65,65,66,67	0
3	SO4	C	409	5/5	0.90	0.15	35,36,42,44	0
4	GOL	B	414	6/6	0.91	0.13	39,45,46,47	0
3	SO4	D	410	5/5	0.93	0.19	59,62,62,63	0
3	SO4	D	408	5/5	0.96	0.08	37,38,41,42	0
3	SO4	A	408	5/5	0.96	0.10	37,38,40,41	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.