



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 17, 2020 – 12:03 PM BST

PDB ID : 6JR7  
Title : Flavobacterium johnsoniae GH31 dextranase, FjDex31A, complexed with glucose  
Authors : Tonozuka, T.  
Deposited on : 2019-04-02  
Resolution : 1.75 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

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

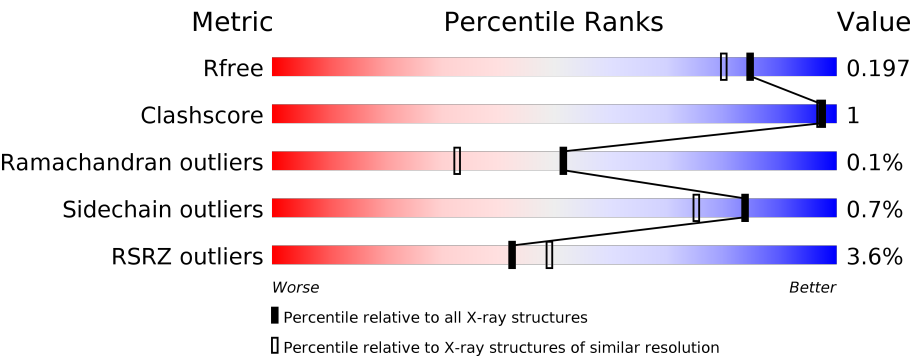
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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.



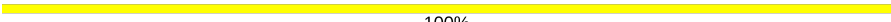


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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

Mol	Chain	Length	Quality of chain
1	A	838	<div><div>4%</div><div><div></div><div>94%</div><div></div><div></div></div><div></div></div>
1	B	838	<div><div>3%</div><div><div></div><div>94%</div><div></div><div></div></div><div></div></div>
1	C	838	<div><div>3%</div><div><div></div><div>95%</div><div></div><div></div></div><div></div></div>
1	D	838	<div><div>4%</div><div><div></div><div>94%</div><div></div><div></div></div><div></div></div>
2	E	2	<div><div></div><div><div></div><div>100%</div><div></div><div></div></div><div></div></div>
2	F	2	<div><div></div><div><div></div><div>50%</div><div></div><div>50%</div></div><div></div></div>

Continued on next page...

*Continued from previous page...*

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

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 29918 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 Candidate alpha-glycosidase Glycoside hydrolase family 31.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	812	Total	C	N	O	S	0	1	0
			6588	4223	1103	1231	31			
1	B	812	Total	C	N	O	S	0	4	0
			6607	4238	1107	1231	31			
1	C	812	Total	C	N	O	S	0	2	0
			6593	4228	1103	1231	31			
1	D	812	Total	C	N	O	S	0	5	0
			6609	4238	1105	1235	31			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	expression tag	UNP A5FBI1
A	0	GLY	-	expression tag	UNP A5FBI1
A	1	SER	-	expression tag	UNP A5FBI1
A	2	SER	-	expression tag	UNP A5FBI1
A	3	HIS	-	expression tag	UNP A5FBI1
A	4	HIS	-	expression tag	UNP A5FBI1
A	5	HIS	-	expression tag	UNP A5FBI1
A	6	HIS	-	expression tag	UNP A5FBI1
A	7	HIS	-	expression tag	UNP A5FBI1
A	8	HIS	-	expression tag	UNP A5FBI1
A	9	SER	-	expression tag	UNP A5FBI1
A	10	SER	-	expression tag	UNP A5FBI1
A	11	GLY	-	expression tag	UNP A5FBI1
A	12	LEU	-	expression tag	UNP A5FBI1
A	13	VAL	-	expression tag	UNP A5FBI1
A	14	PRO	-	expression tag	UNP A5FBI1
A	15	ARG	-	expression tag	UNP A5FBI1
A	16	GLY	-	expression tag	UNP A5FBI1
A	17	SER	-	expression tag	UNP A5FBI1
A	18	HIS	-	expression tag	UNP A5FBI1
A	19	MET	-	expression tag	UNP A5FBI1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	ALA	-	expression tag	UNP A5FBI1
B	-1	MET	-	expression tag	UNP A5FBI1
B	0	GLY	-	expression tag	UNP A5FBI1
B	1	SER	-	expression tag	UNP A5FBI1
B	2	SER	-	expression tag	UNP A5FBI1
B	3	HIS	-	expression tag	UNP A5FBI1
B	4	HIS	-	expression tag	UNP A5FBI1
B	5	HIS	-	expression tag	UNP A5FBI1
B	6	HIS	-	expression tag	UNP A5FBI1
B	7	HIS	-	expression tag	UNP A5FBI1
B	8	HIS	-	expression tag	UNP A5FBI1
B	9	SER	-	expression tag	UNP A5FBI1
B	10	SER	-	expression tag	UNP A5FBI1
B	11	GLY	-	expression tag	UNP A5FBI1
B	12	LEU	-	expression tag	UNP A5FBI1
B	13	VAL	-	expression tag	UNP A5FBI1
B	14	PRO	-	expression tag	UNP A5FBI1
B	15	ARG	-	expression tag	UNP A5FBI1
B	16	GLY	-	expression tag	UNP A5FBI1
B	17	SER	-	expression tag	UNP A5FBI1
B	18	HIS	-	expression tag	UNP A5FBI1
B	19	MET	-	expression tag	UNP A5FBI1
B	20	ALA	-	expression tag	UNP A5FBI1
C	-1	MET	-	expression tag	UNP A5FBI1
C	0	GLY	-	expression tag	UNP A5FBI1
C	1	SER	-	expression tag	UNP A5FBI1
C	2	SER	-	expression tag	UNP A5FBI1
C	3	HIS	-	expression tag	UNP A5FBI1
C	4	HIS	-	expression tag	UNP A5FBI1
C	5	HIS	-	expression tag	UNP A5FBI1
C	6	HIS	-	expression tag	UNP A5FBI1
C	7	HIS	-	expression tag	UNP A5FBI1
C	8	HIS	-	expression tag	UNP A5FBI1
C	9	SER	-	expression tag	UNP A5FBI1
C	10	SER	-	expression tag	UNP A5FBI1
C	11	GLY	-	expression tag	UNP A5FBI1
C	12	LEU	-	expression tag	UNP A5FBI1
C	13	VAL	-	expression tag	UNP A5FBI1
C	14	PRO	-	expression tag	UNP A5FBI1
C	15	ARG	-	expression tag	UNP A5FBI1
C	16	GLY	-	expression tag	UNP A5FBI1
C	17	SER	-	expression tag	UNP A5FBI1

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	18	HIS	-	expression tag	UNP A5FBI1
C	19	MET	-	expression tag	UNP A5FBI1
C	20	ALA	-	expression tag	UNP A5FBI1
D	-1	MET	-	expression tag	UNP A5FBI1
D	0	GLY	-	expression tag	UNP A5FBI1
D	1	SER	-	expression tag	UNP A5FBI1
D	2	SER	-	expression tag	UNP A5FBI1
D	3	HIS	-	expression tag	UNP A5FBI1
D	4	HIS	-	expression tag	UNP A5FBI1
D	5	HIS	-	expression tag	UNP A5FBI1
D	6	HIS	-	expression tag	UNP A5FBI1
D	7	HIS	-	expression tag	UNP A5FBI1
D	8	HIS	-	expression tag	UNP A5FBI1
D	9	SER	-	expression tag	UNP A5FBI1
D	10	SER	-	expression tag	UNP A5FBI1
D	11	GLY	-	expression tag	UNP A5FBI1
D	12	LEU	-	expression tag	UNP A5FBI1
D	13	VAL	-	expression tag	UNP A5FBI1
D	14	PRO	-	expression tag	UNP A5FBI1
D	15	ARG	-	expression tag	UNP A5FBI1
D	16	GLY	-	expression tag	UNP A5FBI1
D	17	SER	-	expression tag	UNP A5FBI1
D	18	HIS	-	expression tag	UNP A5FBI1
D	19	MET	-	expression tag	UNP A5FBI1
D	20	ALA	-	expression tag	UNP A5FBI1

- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose.



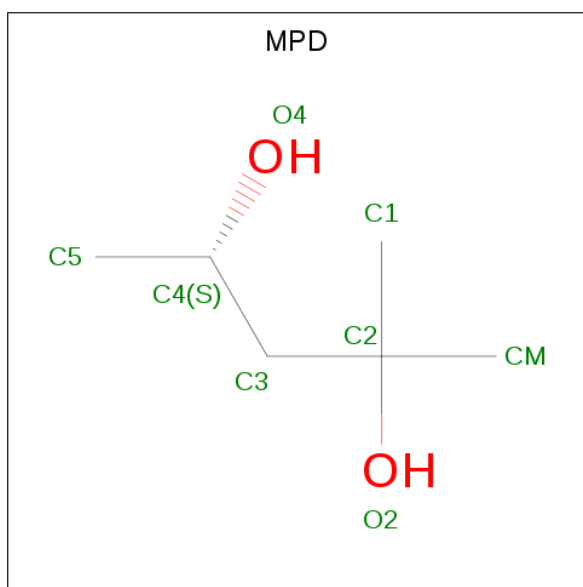
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	E	2	Total	C	O	0	0	0
			23	12	11			
2	F	2	Total	C	O	0	0	0
			23	12	11			
2	G	2	Total	C	O	0	0	0
			23	12	11			
2	H	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



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

- Molecule 4 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula:  $C_6H_{14}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			8	6	2		
4	A	1	Total	C	O	0	0
			8	6	2		
4	B	1	Total	C	O	0	0
			8	6	2		
4	B	1	Total	C	O	0	0
			8	6	2		
4	C	1	Total	C	O	0	0
			8	6	2		
4	C	1	Total	C	O	0	0
			8	6	2		
4	D	1	Total	C	O	0	0
			8	6	2		
4	D	1	Total	C	O	0	0
			8	6	2		

- Molecule 5 is water.

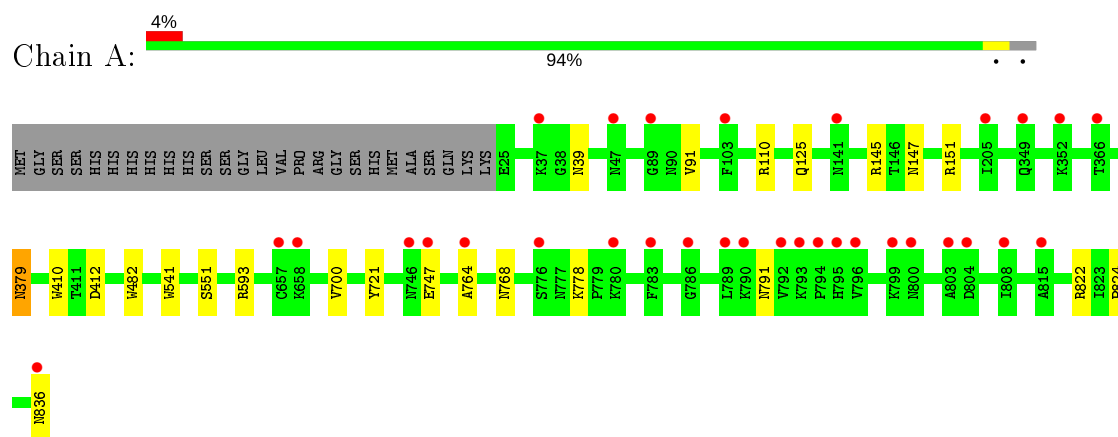
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	817	Total	O	0	0
			817	817		
5	B	870	Total	O	0	0
			870	870		
5	C	830	Total	O	0	0
			830	830		
5	D	828	Total	O	0	0
			828	828		



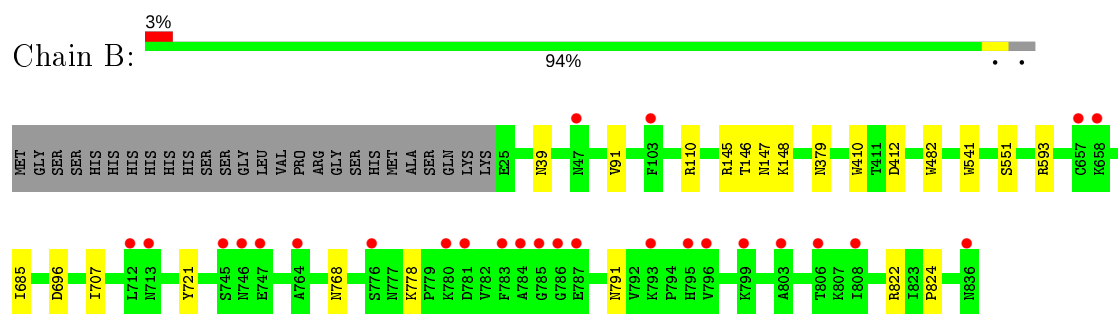
### 3 Residue-property plots [i](#)

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.

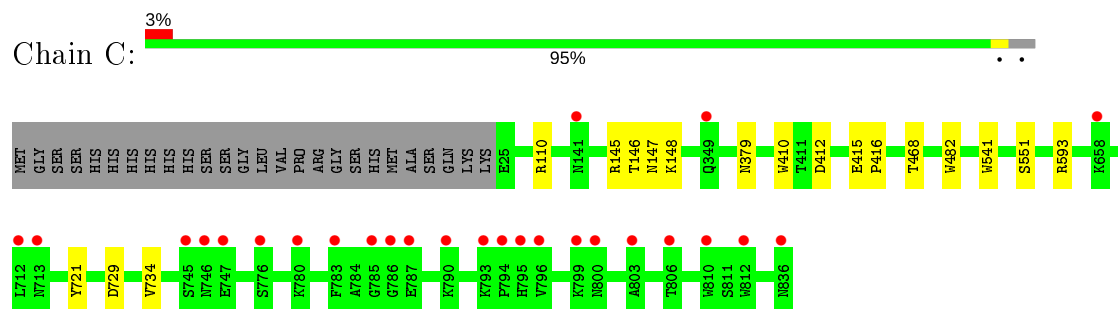
- Molecule 1: Candidate alpha-glycosidase Glycoside hydrolase family 31



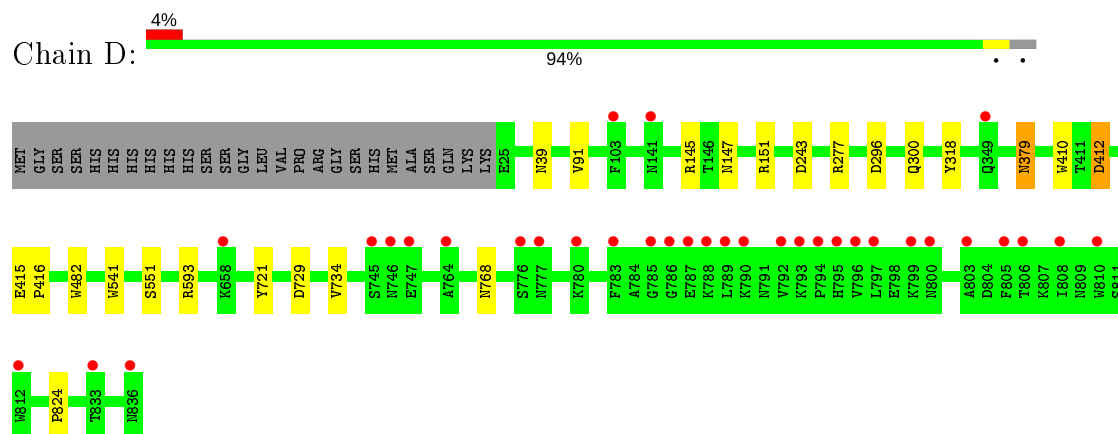
- Molecule 1: Candidate alpha-glycosidase Glycoside hydrolase family 31



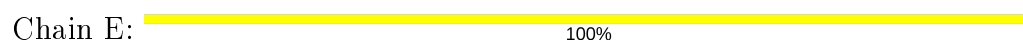
- Molecule 1: Candidate alpha-glycosidase Glycoside hydrolase family 31



- Molecule 1: Candidate alpha-glycosidase Glycoside hydrolase family 31



- Molecule 2: alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose



GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose



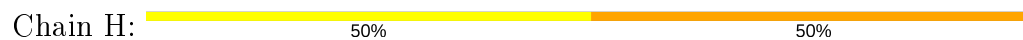
GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose



GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-6)-alpha-D-glucopyranose



GLC1  
GLC2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	116.59Å 102.60Å 114.13Å 87.05° 113.07° 118.53°	Depositor
Resolution (Å)	35.15 – 1.75 35.13 – 1.75	Depositor EDS
% Data completeness (in resolution range)	95.2 (35.15-1.75) 95.2 (35.13-1.75)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.60 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.162 , 0.185 0.176 , 0.197	Depositor DCC
$R_{free}$ test set	20212 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.2	Xtriage
Anisotropy	0.050	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 47.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	29918	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage'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: MPD, GLC, ACT

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	0.44	0/6773	0.66	1/9164 (0.0%)
1	B	0.45	0/6801	0.67	0/9200
1	C	0.45	0/6781	0.68	0/9175
1	D	0.44	0/6806	0.68	3/9208 (0.0%)
All	All	0.44	0/27161	0.67	4/36747 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	243	ASP	CB-CG-OD1	6.39	124.05	118.30
1	A	822	ARG	NE-CZ-NH1	6.04	123.32	120.30
1	D	277	ARG	NE-CZ-NH2	5.58	123.09	120.30
1	D	412	ASP	CB-CG-OD1	5.20	122.97	118.30

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	6588	0	6348	13	0
1	B	6607	0	6385	12	0
1	C	6593	0	6359	10	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	6609	0	6380	11	0
2	E	23	0	21	1	0
2	F	23	0	21	1	0
2	G	23	0	21	1	0
2	H	23	0	21	1	0
3	A	4	0	3	0	0
3	B	4	0	3	0	0
3	C	4	0	3	1	0
3	D	8	0	6	0	0
4	A	16	0	28	0	0
4	B	16	0	28	0	0
4	C	16	0	28	0	0
4	D	16	0	28	0	0
5	A	817	0	0	0	0
5	B	870	0	0	0	0
5	C	830	0	0	0	0
5	D	828	0	0	0	0
All	All	29918	0	25683	43	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 1.

All (43) 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:145:ARG:HE	1:A:147:ASN:HD21	1.28	0.81
1:C:145:ARG:HE	1:C:147:ASN:HD21	1.29	0.79
1:A:379:ASN:HD21	1:B:110:ARG:HH21	1.28	0.77
1:D:145:ARG:HE	1:D:147:ASN:HD21	1.33	0.74
1:C:110:ARG:HH21	1:D:379:ASN:HD21	1.33	0.73
1:B:145:ARG:HE	1:B:147:ASN:HD21	1.40	0.69
1:A:110:ARG:HH21	1:C:379:ASN:HD21	1.42	0.68
1:B:412:ASP:OD2	2:F:1:GLC:H2	2.04	0.57
1:A:412:ASP:OD2	2:E:1:GLC:H2	2.05	0.57
1:D:412:ASP:OD2	2:H:1:GLC:H2	2.05	0.56
1:D:145:ARG:HE	1:D:147:ASN:ND2	2.04	0.53
1:C:145:ARG:HE	1:C:147:ASN:ND2	2.03	0.52
1:A:593:ARG:HD3	1:A:721:TYR:CZ	2.45	0.52
1:B:685[A]:ILE:HD12	1:B:707:ILE:HG12	1.92	0.51
1:A:747:GLU:HA	1:A:836:ASN:HD22	1.77	0.50
1:B:778:LYS:NZ	1:B:791:ASN:HD21	2.10	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:593:ARG:HD3	1:B:721:TYR:CZ	2.47	0.49
1:B:768:ASN:HD22	1:B:824:PRO:HA	1.78	0.49
1:A:39:ASN:HA	1:A:91:VAL:HG23	1.94	0.49
1:B:146:THR:CG2	1:B:148:LYS:HE2	2.43	0.48
1:A:145:ARG:HE	1:A:147:ASN:ND2	2.04	0.48
1:B:145:ARG:HE	1:B:147:ASN:ND2	2.11	0.47
1:D:145:ARG:NE	1:D:147:ASN:HD21	2.09	0.47
1:C:593:ARG:HD3	1:C:721:TYR:CZ	2.51	0.46
1:C:145:ARG:NE	1:C:147:ASN:HD21	2.05	0.46
1:D:415:GLU:N	1:D:416:PRO:HA	2.31	0.46
1:B:685[A]:ILE:CD1	1:B:707:ILE:HG23	2.46	0.45
1:C:412:ASP:OD2	2:G:1:GLC:H2	2.15	0.45
1:C:146:THR:CG2	1:C:148:LYS:HE2	2.47	0.44
1:C:729:ASP:HB3	1:C:734:VAL:HB	1.99	0.44
1:D:729:ASP:HB3	1:D:734:VAL:HB	2.00	0.43
1:D:39:ASN:HA	1:D:91:VAL:HG23	1.98	0.43
1:D:768:ASN:HD22	1:D:824:PRO:HA	1.84	0.43
1:A:700:VAL:HG22	1:A:764:ALA:HB1	2.01	0.43
1:B:39:ASN:HA	1:B:91:VAL:HG23	2.00	0.42
1:D:593:ARG:HD3	1:D:721:TYR:CZ	2.54	0.42
1:C:415:GLU:N	1:C:416:PRO:HA	2.35	0.42
1:A:125:GLN:HG3	3:C:903:ACT:H1	2.02	0.42
1:A:768:ASN:HD22	1:A:824:PRO:HA	1.84	0.42
1:A:145:ARG:NE	1:A:147:ASN:HD21	2.08	0.41
1:B:778:LYS:HZ1	1:B:791:ASN:HD21	1.68	0.41
1:A:778:LYS:NZ	1:A:791:ASN:HD21	2.19	0.41
1:D:300:GLN:NE2	1:D:318:TYR:OH	2.53	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	811/838 (97%)	788 (97%)	21 (3%)	2 (0%)	47	29
1	B	814/838 (97%)	795 (98%)	17 (2%)	2 (0%)	47	29
1	C	812/838 (97%)	788 (97%)	22 (3%)	2 (0%)	47	29
1	D	815/838 (97%)	793 (97%)	20 (2%)	2 (0%)	47	29
All	All	3252/3352 (97%)	3164 (97%)	80 (2%)	8 (0%)	51	29

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	551[A]	SER
1	A	551[B]	SER
1	B	551[A]	SER
1	B	551[B]	SER
1	C	551[A]	SER
1	C	551[B]	SER
1	D	551[A]	SER
1	D	551[B]	SER

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	704/725 (97%)	699 (99%)	5 (1%)	84	75
1	B	707/725 (98%)	701 (99%)	6 (1%)	81	72
1	C	705/725 (97%)	701 (99%)	4 (1%)	86	79
1	D	708/725 (98%)	702 (99%)	6 (1%)	81	72
All	All	2824/2900 (97%)	2803 (99%)	21 (1%)	84	75

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	151	ARG
1	A	379	ASN
1	A	410	TRP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	482	TRP
1	A	541	TRP
1	B	379	ASN
1	B	410	TRP
1	B	482	TRP
1	B	541	TRP
1	B	696	ASP
1	B	822	ARG
1	C	410	TRP
1	C	468	THR
1	C	482	TRP
1	C	541	TRP
1	D	151	ARG
1	D	296	ASP
1	D	379	ASN
1	D	410	TRP
1	D	482	TRP
1	D	541	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (42) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	112	ASN
1	A	125	GLN
1	A	147	ASN
1	A	250	ASN
1	A	300	GLN
1	A	379	ASN
1	A	501	ASN
1	A	768	ASN
1	A	791	ASN
1	A	800	ASN
1	A	829	ASN
1	A	836	ASN
1	B	112	ASN
1	B	125	GLN
1	B	147	ASN
1	B	250	ASN
1	B	300	GLN
1	B	348	GLN
1	B	379	ASN
1	B	501	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	655	ASN
1	B	768	ASN
1	B	791	ASN
1	C	112	ASN
1	C	147	ASN
1	C	250	ASN
1	C	300	GLN
1	C	379	ASN
1	C	501	ASN
1	C	655	ASN
1	C	768	ASN
1	C	817	ASN
1	D	47	ASN
1	D	112	ASN
1	D	147	ASN
1	D	250	ASN
1	D	300	GLN
1	D	319	ASN
1	D	379	ASN
1	D	655	ASN
1	D	768	ASN
1	D	829	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

8 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GLC	E	1	2	12,12,12	0.58	0	17,17,17	0.74	0
2	GLC	E	2	2	11,11,12	0.61	0	15,15,17	1.14	3 (20%)
2	GLC	F	1	2	12,12,12	0.52	0	17,17,17	1.00	1 (5%)
2	GLC	F	2	2	11,11,12	0.74	0	15,15,17	1.45	4 (26%)
2	GLC	G	1	2	12,12,12	0.66	0	17,17,17	0.73	0
2	GLC	G	2	2	11,11,12	0.74	0	15,15,17	1.10	1 (6%)
2	GLC	H	1	2	12,12,12	0.65	0	17,17,17	0.87	1 (5%)
2	GLC	H	2	2	11,11,12	0.90	0	15,15,17	1.05	1 (6%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GLC	E	1	2	-	2/2/22/22	0/1/1/1
2	GLC	E	2	2	-	0/2/19/22	0/1/1/1
2	GLC	F	1	2	-	2/2/22/22	0/1/1/1
2	GLC	F	2	2	-	0/2/19/22	0/1/1/1
2	GLC	G	1	2	-	2/2/22/22	0/1/1/1
2	GLC	G	2	2	-	0/2/19/22	0/1/1/1
2	GLC	H	1	2	-	2/2/22/22	0/1/1/1
2	GLC	H	2	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	2	GLC	O5-C5-C6	2.67	111.38	107.20
2	G	2	GLC	O5-C5-C6	2.57	111.23	107.20
2	E	2	GLC	C1-C2-C3	2.57	112.82	109.67
2	F	2	GLC	C1-O5-C5	2.56	115.66	112.19
2	F	2	GLC	C1-C2-C3	2.49	112.73	109.67
2	E	2	GLC	C1-O5-C5	2.17	115.13	112.19
2	F	2	GLC	O5-C1-C2	2.14	114.07	110.77
2	H	2	GLC	O5-C5-C6	2.10	110.50	107.20
2	H	1	GLC	O2-C2-C3	2.03	115.03	110.35
2	F	1	GLC	O2-C2-C3	2.01	115.01	110.35
2	E	2	GLC	O5-C1-C2	2.00	113.86	110.77

There are no chirality outliers.

All (8) torsion outliers are listed below:

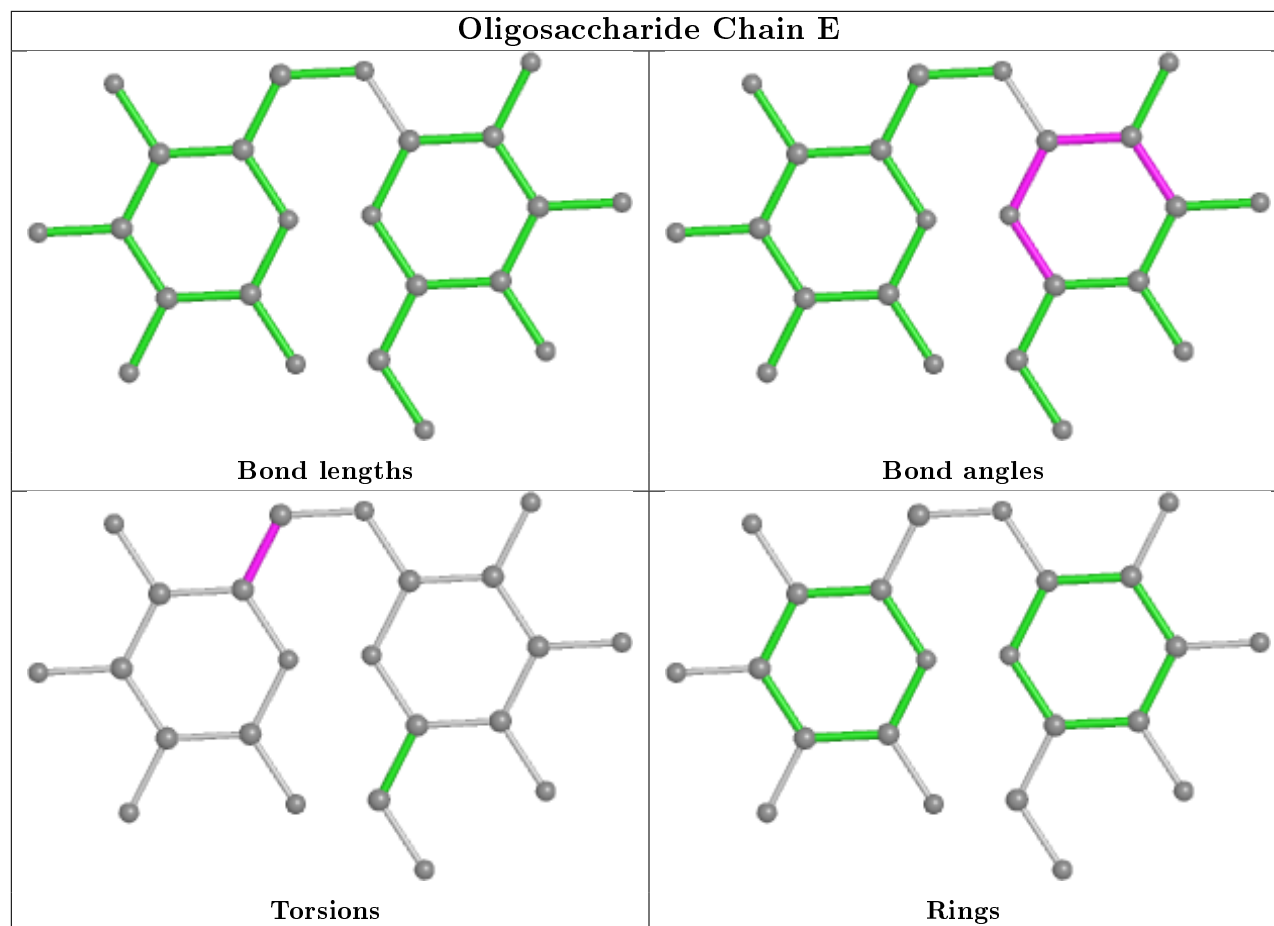
Mol	Chain	Res	Type	Atoms
2	G	1	GLC	O5-C5-C6-O6
2	F	1	GLC	O5-C5-C6-O6
2	E	1	GLC	O5-C5-C6-O6
2	H	1	GLC	O5-C5-C6-O6
2	E	1	GLC	C4-C5-C6-O6
2	F	1	GLC	C4-C5-C6-O6
2	H	1	GLC	C4-C5-C6-O6
2	G	1	GLC	C4-C5-C6-O6

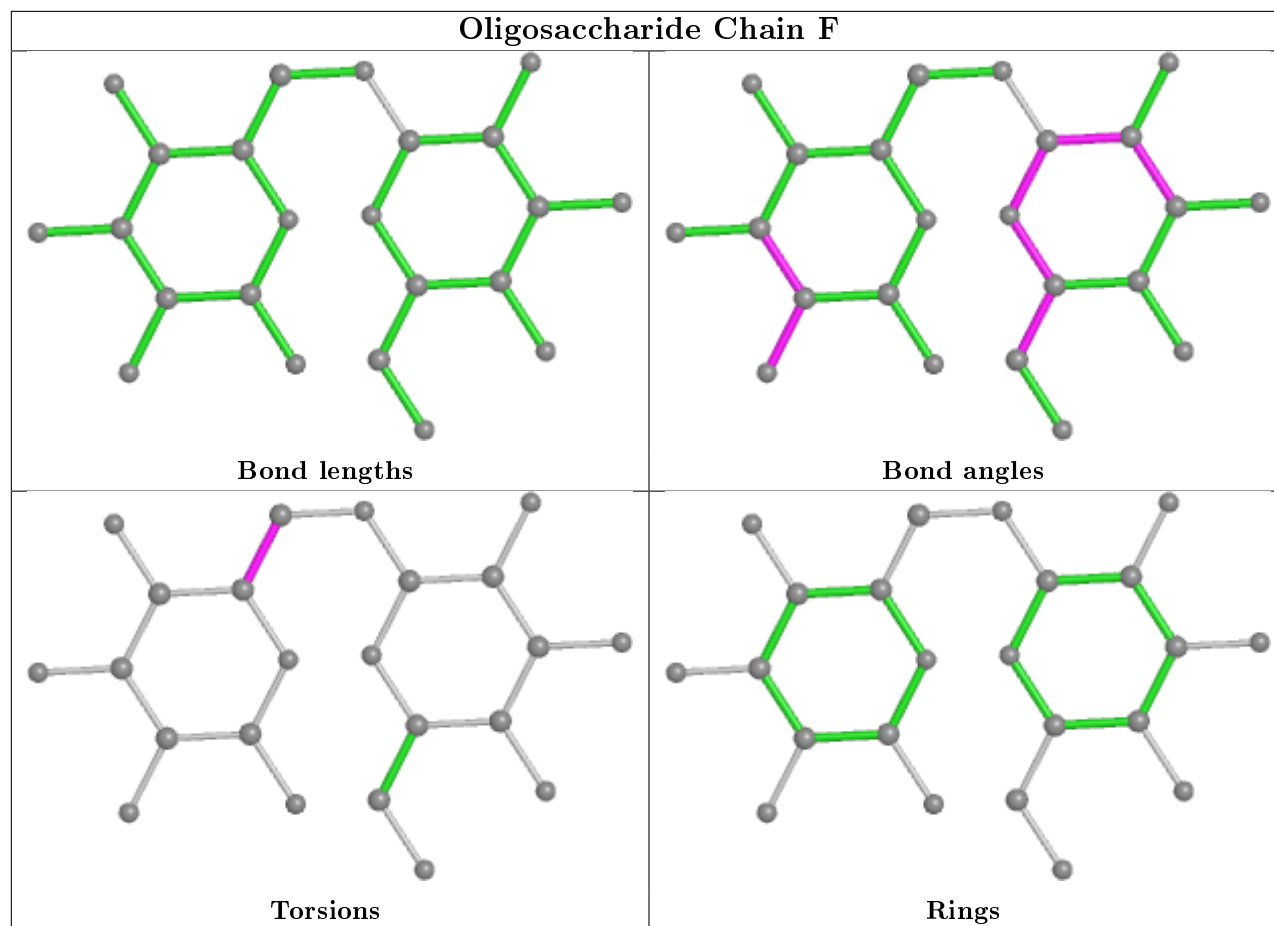
There are no ring outliers.

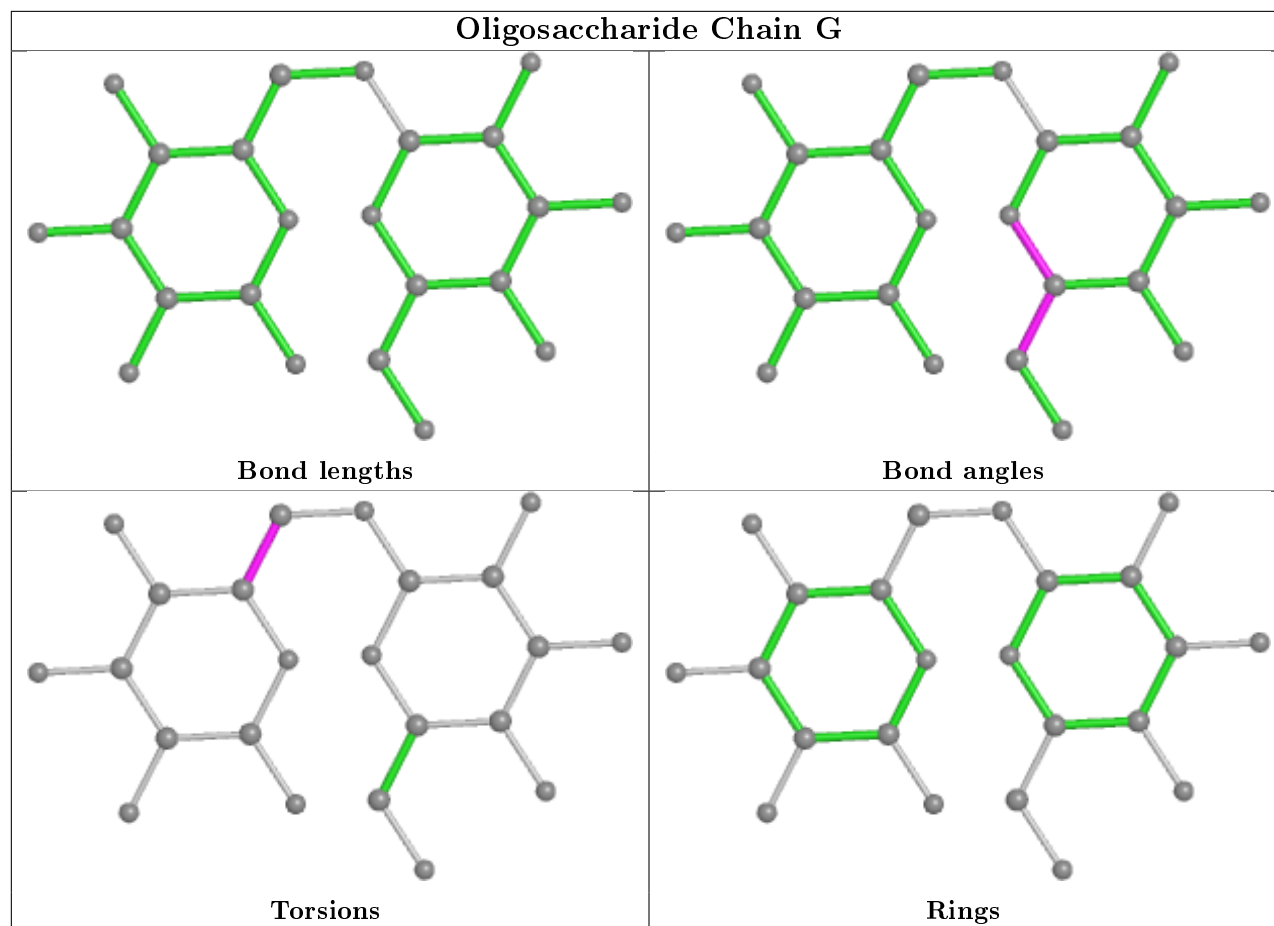
4 monomers are involved in 4 short contacts:

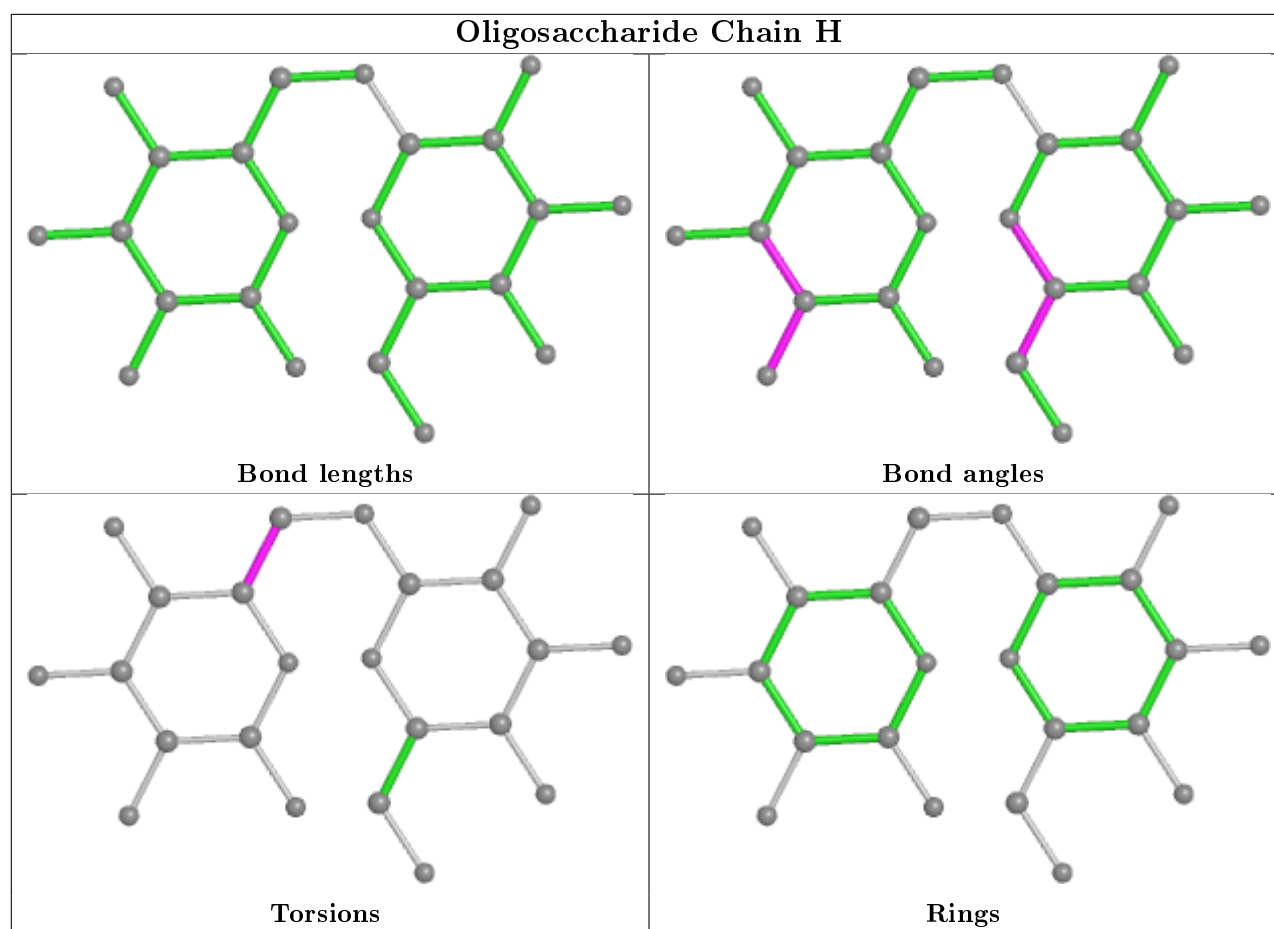
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	1	GLC	1	0
2	E	1	GLC	1	0
2	F	1	GLC	1	0
2	H	1	GLC	1	0

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](#)

13 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
4	MPD	C	905	-	7,7,7	0.27	0	9,10,10	0.29	0
4	MPD	A	904	-	7,7,7	0.27	0	9,10,10	0.34	0
4	MPD	B	905	-	7,7,7	0.22	0	9,10,10	0.27	0
4	MPD	D	906	-	7,7,7	0.27	0	9,10,10	0.24	0
4	MPD	A	905	-	7,7,7	0.27	0	9,10,10	0.29	0
3	ACT	C	903	-	1,3,3	0.92	0	0,3,3	0.00	-
3	ACT	A	903	-	1,3,3	1.68	0	0,3,3	0.00	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	MPD	C	904	-	7,7,7	0.36	0	9,10,10	0.30	0
3	ACT	D	904	-	1,3,3	1.21	0	0,3,3	0.00	-
3	ACT	D	903	-	1,3,3	1.71	0	0,3,3	0.00	-
3	ACT	B	903	-	1,3,3	1.33	0	0,3,3	0.00	-
4	MPD	B	904	-	7,7,7	0.28	0	9,10,10	0.36	0
4	MPD	D	905	-	7,7,7	0.26	0	9,10,10	0.29	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	MPD	C	905	-	-	0/5/5/5	-
4	MPD	B	905	-	-	0/5/5/5	-
4	MPD	D	906	-	-	0/5/5/5	-
4	MPD	A	905	-	-	0/5/5/5	-
4	MPD	A	904	-	-	0/5/5/5	-
4	MPD	C	904	-	-	0/5/5/5	-
4	MPD	B	904	-	-	0/5/5/5	-
4	MPD	D	905	-	-	0/5/5/5	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	903	ACT	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

### 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	812/838 (96%)	0.10	32 (3%) 39 45	14, 22, 42, 73	0
1	B	812/838 (96%)	0.09	26 (3%) 47 54	13, 20, 46, 72	0
1	C	812/838 (96%)	0.11	26 (3%) 47 54	14, 21, 48, 81	0
1	D	812/838 (96%)	0.09	34 (4%) 36 42	14, 21, 47, 74	0
All	All	3248/3352 (96%)	0.10	118 (3%) 42 49	13, 21, 46, 81	0

All (118) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	786	GLY	6.8
1	C	747	GLU	6.8
1	D	747	GLU	6.4
1	D	795	HIS	5.8
1	B	836	ASN	5.6
1	A	747	GLU	5.6
1	C	796	VAL	5.5
1	C	712	LEU	5.2
1	C	836	ASN	5.1
1	C	786	GLY	5.0
1	B	764	ALA	4.9
1	B	795	HIS	4.6
1	B	747	GLU	4.6
1	B	746	ASN	4.6
1	D	658	LYS	4.4
1	C	746	ASN	4.4
1	C	658	LYS	4.3
1	A	795	HIS	4.3
1	A	803	ALA	4.3
1	A	796	VAL	4.2
1	C	780	LYS	4.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	796	VAL	4.2
1	A	103	PHE	4.0
1	C	783	PHE	4.0
1	C	795	HIS	4.0
1	B	783	PHE	3.7
1	D	786	GLY	3.6
1	A	658	LYS	3.5
1	C	799	LYS	3.5
1	D	803	ALA	3.5
1	B	103	PHE	3.5
1	C	806	THR	3.5
1	C	800	ASN	3.4
1	A	786	GLY	3.4
1	D	793	LYS	3.4
1	D	746	ASN	3.4
1	B	787	GLU	3.4
1	C	745	SER	3.3
1	D	805	PHE	3.3
1	C	810	TRP	3.3
1	B	745	SER	3.3
1	D	799	LYS	3.3
1	B	793	LYS	3.2
1	C	776	SER	3.2
1	A	776	SER	3.2
1	C	793	LYS	3.1
1	C	785	GLY	3.1
1	D	783	PHE	3.1
1	D	836	ASN	3.1
1	B	658	LYS	3.1
1	D	792	VAL	3.1
1	D	806	THR	3.0
1	B	780	LYS	3.0
1	B	806	THR	3.0
1	C	790	LYS	2.9
1	D	787	GLU	2.9
1	B	808	ILE	2.9
1	D	808	ILE	2.9
1	D	776	SER	2.9
1	B	712	LEU	2.8
1	A	349	GLN	2.8
1	A	746	ASN	2.8
1	A	836	ASN	2.8

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	799	LYS	2.8
1	D	797	LEU	2.8
1	D	780	LYS	2.8
1	C	713	ASN	2.7
1	A	141	ASN	2.7
1	D	800	ASN	2.7
1	D	790	LYS	2.7
1	B	776	SER	2.6
1	C	787	GLU	2.6
1	A	815	ALA	2.6
1	C	812	TRP	2.6
1	D	794	PRO	2.6
1	C	803	ALA	2.5
1	D	777	ASN	2.5
1	A	793	LYS	2.5
1	D	745	SER	2.5
1	A	89	GLY	2.5
1	A	352	LYS	2.4
1	D	103	PHE	2.4
1	A	800	ASN	2.4
1	B	47	ASN	2.4
1	D	788	LYS	2.4
1	A	764	ALA	2.4
1	B	785	GLY	2.4
1	A	657	CYS	2.3
1	A	808	ILE	2.3
1	A	780	LYS	2.3
1	A	790	LYS	2.3
1	A	783	PHE	2.3
1	A	37	LYS	2.3
1	D	812	TRP	2.3
1	B	784	ALA	2.3
1	C	141	ASN	2.3
1	C	349	GLN	2.2
1	B	781	ASP	2.2
1	A	47	ASN	2.2
1	A	366	THR	2.2
1	D	789	LEU	2.2
1	A	799	LYS	2.2
1	D	764	ALA	2.2
1	D	349	GLN	2.1
1	D	833	THR	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	810	TRP	2.1
1	A	789	LEU	2.1
1	A	792	VAL	2.1
1	A	804	ASP	2.1
1	A	205	ILE	2.1
1	A	794	PRO	2.1
1	C	794	PRO	2.1
1	B	657	CYS	2.0
1	D	785	GLY	2.0
1	B	803	ALA	2.0
1	B	796	VAL	2.0
1	B	713	ASN	2.0
1	D	141	ASN	2.0

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

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

## 6.3 Carbohydrates ⓘ

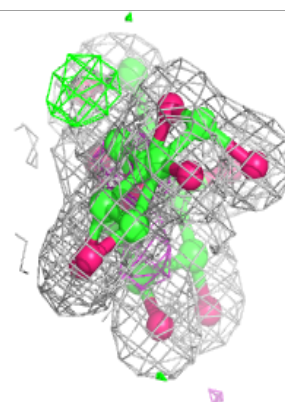
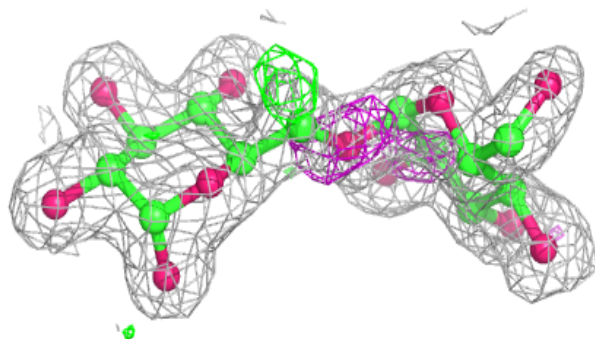
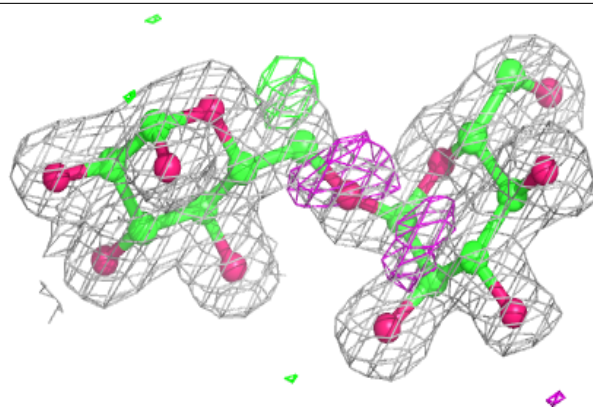
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<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	F	2	11/12	0.70	0.28	27,41,45,45	0
2	GLC	H	2	11/12	0.70	0.28	29,40,43,44	0
2	GLC	G	2	11/12	0.77	0.22	32,40,42,43	0
2	GLC	H	1	12/12	0.86	0.15	26,29,33,36	0
2	GLC	E	2	11/12	0.90	0.16	28,33,36,36	0
2	GLC	G	1	12/12	0.91	0.12	24,25,30,35	0
2	GLC	F	1	12/12	0.91	0.13	20,26,30,36	0
2	GLC	E	1	12/12	0.94	0.11	24,27,30,32	0

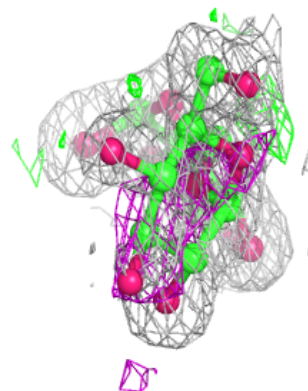
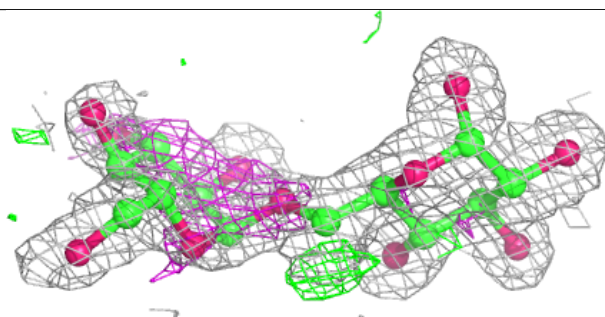
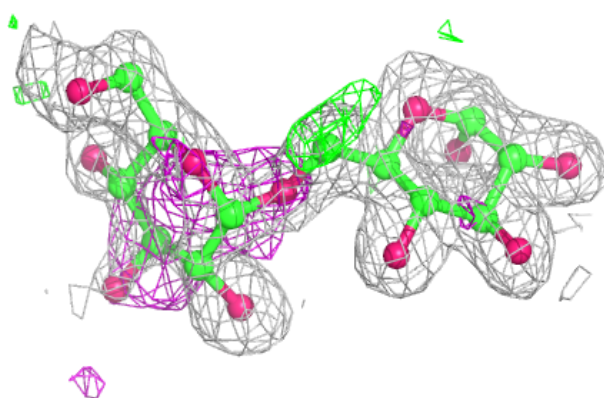
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**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)

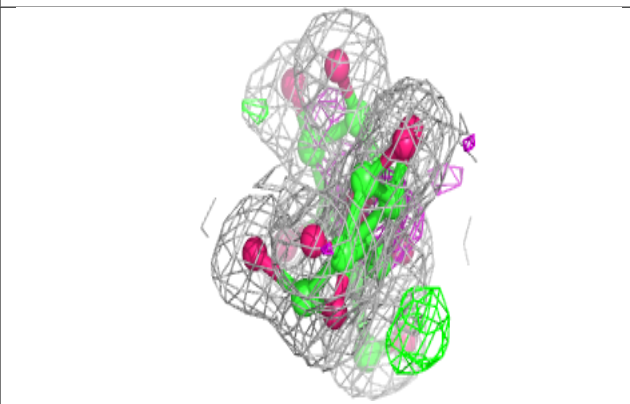
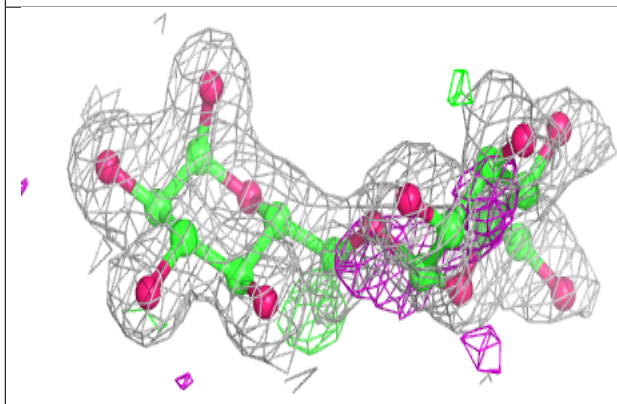
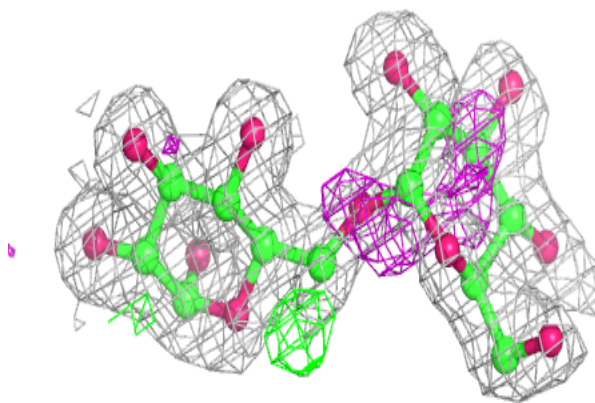
**Electron density around Chain F:**

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

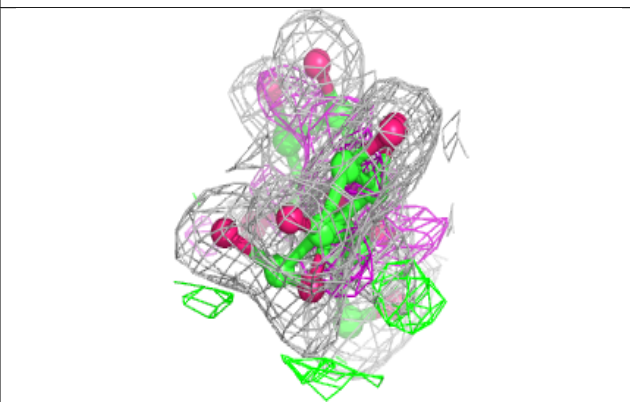
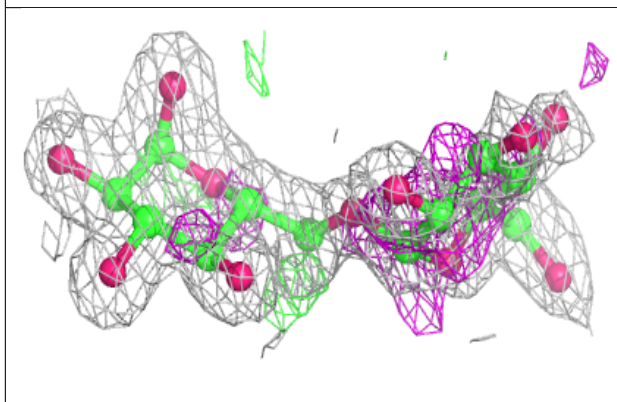
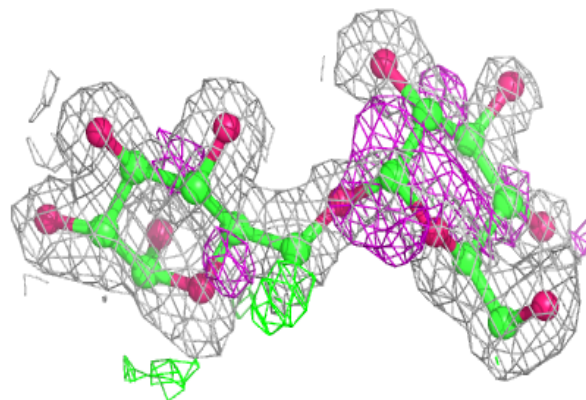


**Electron density around Chain G:**

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

**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)



## 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	MPD	C	905	8/8	0.68	0.28	68,70,72,73	0
4	MPD	A	905	8/8	0.72	0.24	59,63,65,66	0
4	MPD	D	906	8/8	0.81	0.27	62,63,65,67	0
4	MPD	B	905	8/8	0.88	0.13	39,40,44,47	0
4	MPD	A	904	8/8	0.89	0.16	42,43,45,49	0
4	MPD	D	905	8/8	0.89	0.16	36,39,39,39	0
3	ACT	B	903	4/4	0.90	0.12	29,29,30,31	0
3	ACT	D	904	4/4	0.91	0.15	27,29,29,30	0
4	MPD	C	904	8/8	0.92	0.12	37,38,39,39	0
4	MPD	B	904	8/8	0.93	0.12	32,33,34,35	0
3	ACT	A	903	4/4	0.93	0.13	24,25,25,25	0
3	ACT	D	903	4/4	0.94	0.12	27,29,29,30	0
3	ACT	C	903	4/4	0.96	0.11	24,24,25,25	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.