



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 21, 2020 – 07:16 AM BST

PDB ID : 6S21  
Title : Metabolism of multiple glycosaminoglycans by bacteroides thetaiotaomicron is orchestrated by a versatile core genetic locus (BT33494S-sulf)  
Authors : Ndeh, D.; Basle, A.; Strahl, H.; Henrissat, B.; Terrapon, N.; Cartmell, A.  
Deposited on : 2019-06-19  
Resolution : 2.80 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

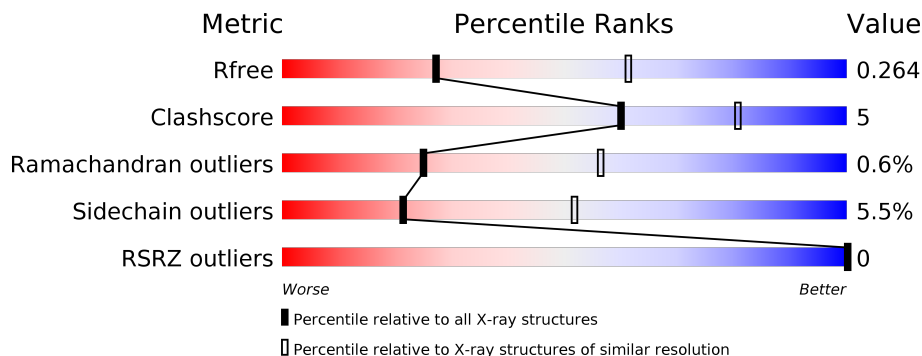
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	513	 79% 15% • 5%
1	B	513	 81% 12% • 5%
1	C	513	 80% 13% • 5%
2	D	3	 33% 67%
2	E	3	 100%
2	F	3	 33% 67%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11684 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 Endo-4-O-sulfatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	486	3841	2434	648	740	19	0	0	0
1	B	485	3824	2424	642	739	19	0	0	0
1	C	486	3841	2435	650	737	19	0	0	0

There are 63 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	MET	-	initiating methionine	UNP Q8A2F6
A	-3	GLY	-	expression tag	UNP Q8A2F6
A	-2	SER	-	expression tag	UNP Q8A2F6
A	-1	SER	-	expression tag	UNP Q8A2F6
A	0	HIS	-	expression tag	UNP Q8A2F6
A	1	HIS	-	expression tag	UNP Q8A2F6
A	2	HIS	-	expression tag	UNP Q8A2F6
A	3	HIS	-	expression tag	UNP Q8A2F6
A	4	HIS	-	expression tag	UNP Q8A2F6
A	5	HIS	-	expression tag	UNP Q8A2F6
A	6	SER	-	expression tag	UNP Q8A2F6
A	7	SER	-	expression tag	UNP Q8A2F6
A	8	GLY	-	expression tag	UNP Q8A2F6
A	9	LEU	-	expression tag	UNP Q8A2F6
A	10	VAL	-	expression tag	UNP Q8A2F6
A	11	PRO	-	expression tag	UNP Q8A2F6
A	12	ARG	-	expression tag	UNP Q8A2F6
A	13	GLY	-	expression tag	UNP Q8A2F6
A	14	SER	-	expression tag	UNP Q8A2F6
A	15	HIS	-	expression tag	UNP Q8A2F6
A	16	MET	-	expression tag	UNP Q8A2F6
B	-4	MET	-	initiating methionine	UNP Q8A2F6
B	-3	GLY	-	expression tag	UNP Q8A2F6

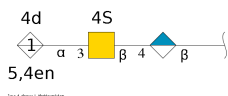
*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	SER	-	expression tag	UNP Q8A2F6
B	-1	SER	-	expression tag	UNP Q8A2F6
B	0	HIS	-	expression tag	UNP Q8A2F6
B	1	HIS	-	expression tag	UNP Q8A2F6
B	2	HIS	-	expression tag	UNP Q8A2F6
B	3	HIS	-	expression tag	UNP Q8A2F6
B	4	HIS	-	expression tag	UNP Q8A2F6
B	5	HIS	-	expression tag	UNP Q8A2F6
B	6	SER	-	expression tag	UNP Q8A2F6
B	7	SER	-	expression tag	UNP Q8A2F6
B	8	GLY	-	expression tag	UNP Q8A2F6
B	9	LEU	-	expression tag	UNP Q8A2F6
B	10	VAL	-	expression tag	UNP Q8A2F6
B	11	PRO	-	expression tag	UNP Q8A2F6
B	12	ARG	-	expression tag	UNP Q8A2F6
B	13	GLY	-	expression tag	UNP Q8A2F6
B	14	SER	-	expression tag	UNP Q8A2F6
B	15	HIS	-	expression tag	UNP Q8A2F6
B	16	MET	-	expression tag	UNP Q8A2F6
C	-4	MET	-	initiating methionine	UNP Q8A2F6
C	-3	GLY	-	expression tag	UNP Q8A2F6
C	-2	SER	-	expression tag	UNP Q8A2F6
C	-1	SER	-	expression tag	UNP Q8A2F6
C	0	HIS	-	expression tag	UNP Q8A2F6
C	1	HIS	-	expression tag	UNP Q8A2F6
C	2	HIS	-	expression tag	UNP Q8A2F6
C	3	HIS	-	expression tag	UNP Q8A2F6
C	4	HIS	-	expression tag	UNP Q8A2F6
C	5	HIS	-	expression tag	UNP Q8A2F6
C	6	SER	-	expression tag	UNP Q8A2F6
C	7	SER	-	expression tag	UNP Q8A2F6
C	8	GLY	-	expression tag	UNP Q8A2F6
C	9	LEU	-	expression tag	UNP Q8A2F6
C	10	VAL	-	expression tag	UNP Q8A2F6
C	11	PRO	-	expression tag	UNP Q8A2F6
C	12	ARG	-	expression tag	UNP Q8A2F6
C	13	GLY	-	expression tag	UNP Q8A2F6
C	14	SER	-	expression tag	UNP Q8A2F6
C	15	HIS	-	expression tag	UNP Q8A2F6
C	16	MET	-	expression tag	UNP Q8A2F6

- Molecule 2 is an oligosaccharide called 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-3)-2-acetamido-2-deoxy-4-O-sulfo-beta-D-galactopyranose-(1-4)-beta-D-glucopyranuronic

acid.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	3	Total	C	N	O	S	0	0	0
			42	20	1	20	1			
2	E	3	Total	C	N	O	S	0	0	0
			42	20	1	20	1			
2	F	3	Total	C	N	O	S	0	0	0
			42	20	1	20	1			

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Ca	0	0
			1	1		
3	A	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		


- Molecule 4 is water.

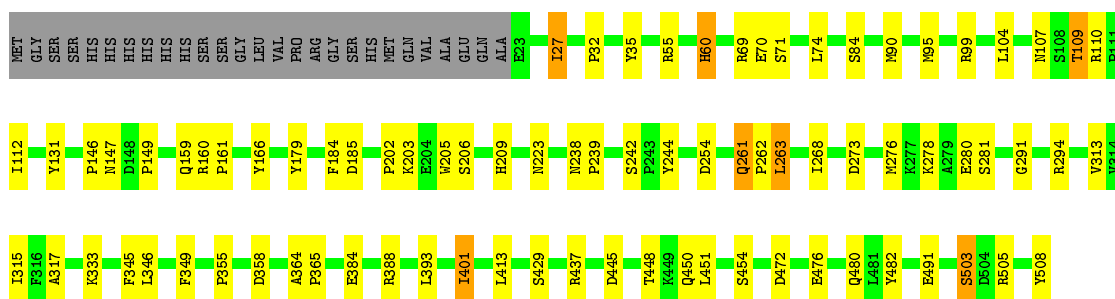
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	16	Total	O	0	0
			16	16		
4	B	16	Total	O	0	0
			16	16		
4	C	17	Total	O	0	0
			17	17		

### 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.

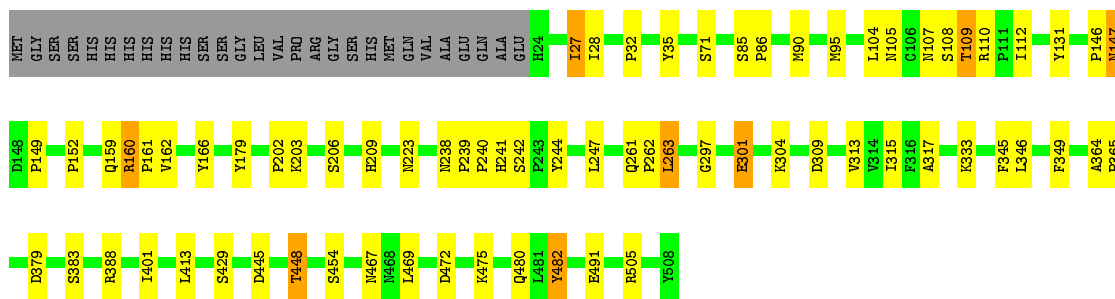
- Molecule 1: Endo-4-O-sulfatase

Chain A: 




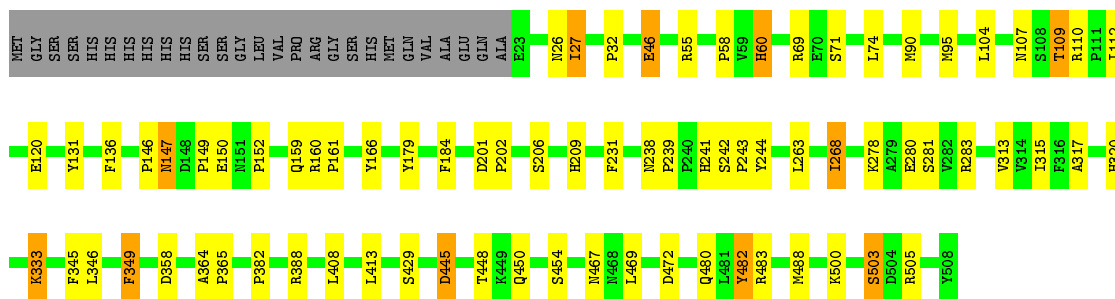
- Molecule 1: Endo-4-O-sulfatase

Chain B: 



- Molecule 1: Endo-4-O-sulfatase

Chain C: 

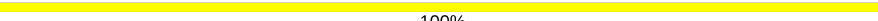


- Molecule 2: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-3)-2-acetamido-2-deoxy-4-O-sulfo-beta-D-galactopyranose-(1-4)-beta-D-glucopyranuronic acid

Chain D:  33% 67%

BDP1  
ASG2  
GCD3

- Molecule 2: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-3)-2-acetamido-2-deoxy-4-O-sulfo-beta-D-galactopyranose-(1-4)-beta-D-glucopyranuronic acid

Chain E:  100%

BDP1  
ASG2  
GCD3

- Molecule 2: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-3)-2-acetamido-2-deoxy-4-O-sulfo-beta-D-galactopyranose-(1-4)-beta-D-glucopyranuronic acid

Chain F:  33% 67%

BDP1  
ASG2  
GCD3

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.91Å 93.91Å 304.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	68.98 – 2.80 68.89 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (68.98-2.80) 100.0 (68.89-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.65 (at 2.81Å)	Xtrriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.189 , 0.261 0.198 , 0.264	Depositor DCC
$R_{free}$ test set	1856 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.3	Xtrriage
Anisotropy	0.264	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 37.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	11684	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.39% 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: CA, ASG, GCD, BDP

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	0.83	2/3950 (0.1%)	1.06	4/5384 (0.1%)
1	B	0.82	0/3933	1.03	3/5364 (0.1%)
1	C	0.84	2/3950 (0.1%)	1.07	4/5383 (0.1%)
All	All	0.83	4/11833 (0.0%)	1.05	11/16131 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3
1	B	0	3
1	C	0	3
All	All	0	9

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	46	GLU	CD-OE2	5.58	1.31	1.25
1	C	150	GLU	CD-OE1	5.37	1.31	1.25
1	A	70	GLU	CD-OE2	-5.16	1.20	1.25
1	A	355	PRO	C-O	-5.04	1.13	1.23

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	60	HIS	CB-CA-C	6.69	123.78	110.40
1	A	508	TYR	CA-C-O	6.24	133.20	120.10
1	C	482	TYR	CB-CG-CD1	5.95	124.57	121.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	60	HIS	CB-CA-C	5.89	122.17	110.40
1	A	69	ARG	NE-CZ-NH1	5.73	123.16	120.30
1	B	475	LYS	CB-CA-C	5.37	121.13	110.40
1	B	203	LYS	CB-CA-C	5.28	120.97	110.40
1	B	482	TYR	CB-CG-CD1	5.12	124.07	121.00
1	A	203	LYS	CB-CA-C	5.12	120.64	110.40
1	C	69	ARG	NE-CZ-NH1	5.02	122.81	120.30
1	C	152	PRO	N-CA-C	5.01	125.12	112.10

There are no chirality outliers.

All (9) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	131	TYR	Peptide
1	A	146	PRO	Peptide
1	A	147	ASN	Peptide
1	B	131	TYR	Peptide
1	B	146	PRO	Peptide
1	B	147	ASN	Peptide
1	C	131	TYR	Peptide
1	C	146	PRO	Peptide
1	C	147	ASN	Peptide

## 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	3841	0	3585	33	1
1	B	3824	0	3558	34	0
1	C	3841	0	3594	36	1
2	D	42	0	20	2	0
2	E	42	0	20	0	0
2	F	42	0	20	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	16	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	16	0	0	4	0
4	C	17	0	0	2	0
All	All	11684	0	10797	104	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All (104) 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:483:ARG:CB	4:C:716:HOH:O	2.16	0.93
1:A:185:ASP:OD2	2:D:2:ASG:O6	2.00	0.79
1:A:107:ASN:OD1	1:A:109:THR:HB	1.86	0.76
1:A:242:SER:HB2	2:D:1:BDP:H1	1.69	0.75
1:B:107:ASN:OD1	1:B:109:THR:HB	1.87	0.75
1:B:147:ASN:ND2	1:B:166:TYR:O	2.21	0.73
1:C:107:ASN:OD1	1:C:109:THR:HB	1.88	0.73
1:C:147:ASN:ND2	1:C:166:TYR:O	2.22	0.72
1:A:451:LEU:O	1:B:247:LEU:HD12	1.90	0.70
1:C:238:ASN:HB3	1:C:239:PRO:HD3	1.74	0.70
1:A:238:ASN:HB3	1:A:239:PRO:HD3	1.74	0.70
1:B:238:ASN:HB3	1:B:239:PRO:HD3	1.74	0.70
1:B:179:TYR:HH	1:B:209:HIS:HE2	1.41	0.69
1:B:297:GLY:O	1:B:301:GLU:HG2	1.95	0.67
1:C:120:GLU:HG2	1:C:382:PRO:HG3	1.78	0.65
1:A:109:THR:HG22	1:A:110:ARG:HG3	1.82	0.62
1:C:448:THR:O	1:C:450:GLN:N	2.32	0.62
1:A:202:PRO:HG3	1:A:209:HIS:CE1	2.35	0.61
1:B:346:LEU:HD12	1:B:346:LEU:N	2.17	0.60
1:B:202:PRO:HG3	1:B:209:HIS:CE1	2.37	0.59
1:A:448:THR:O	1:A:450:GLN:N	2.33	0.59
1:B:179:TYR:OH	1:B:209:HIS:NE2	2.30	0.58
1:B:109:THR:HG22	1:B:110:ARG:HG3	1.86	0.58
1:B:304:LYS:HE3	1:B:309:ASP:OD2	2.03	0.58
1:C:346:LEU:N	1:C:346:LEU:HD12	2.19	0.57
1:C:467:ASN:O	1:C:469:LEU:HD13	2.04	0.57
1:B:107:ASN:ND2	4:B:701:HOH:O	2.38	0.56
1:A:55:ARG:HB3	1:A:268:ILE:HD12	1.88	0.56
1:C:109:THR:HG22	1:C:110:ARG:HG3	1.87	0.55
1:C:243:PRO:O	4:C:701:HOH:O	2.18	0.55
1:C:202:PRO:HG3	1:C:209:HIS:CE1	2.42	0.55

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:280:GLU:OE2	1:C:283:ARG:NH1	2.39	0.54
1:A:346:LEU:N	1:A:346:LEU:HD12	2.22	0.54
1:C:364:ALA:HB3	1:C:365:PRO:HD3	1.90	0.54
1:C:58:PRO:HD3	1:C:268:ILE:HD13	1.90	0.54
1:B:345:PHE:C	1:B:346:LEU:HD12	2.29	0.53
1:C:345:PHE:C	1:C:346:LEU:HD12	2.29	0.52
1:C:448:THR:HG23	1:C:448:THR:O	2.10	0.52
1:A:345:PHE:C	1:A:346:LEU:HD12	2.30	0.52
1:B:364:ALA:HB3	1:B:365:PRO:HD3	1.92	0.51
1:C:445:ASP:OD1	1:C:448:THR:HG22	2.11	0.51
1:A:445:ASP:OD1	1:A:448:THR:HG22	2.11	0.50
1:A:364:ALA:HB3	1:A:365:PRO:HD3	1.94	0.50
1:B:90:MET:HG2	1:B:95:MET:O	2.12	0.49
1:A:448:THR:O	1:A:448:THR:HG23	2.13	0.49
1:B:105:ASN:HB3	4:B:701:HOH:O	2.14	0.48
1:A:90:MET:HG2	1:A:95:MET:O	2.14	0.48
1:C:278:LYS:O	1:C:281:SER:HB2	2.13	0.48
1:A:27:ILE:HA	1:A:313:VAL:O	2.14	0.48
1:B:160:ARG:HA	1:B:161:PRO:C	2.34	0.48
1:B:152:PRO:CG	4:B:713:HOH:O	2.62	0.48
1:A:254:ASP:OD1	1:A:294:ARG:NH1	2.31	0.48
1:B:315:ILE:HA	1:B:346:LEU:O	2.14	0.48
1:B:445:ASP:OD2	1:B:448:THR:HG22	2.13	0.48
1:C:90:MET:HG2	1:C:95:MET:O	2.14	0.48
1:A:315:ILE:HA	1:A:346:LEU:O	2.15	0.47
1:A:413:LEU:HD12	1:A:429:SER:HB3	1.97	0.47
1:A:313:VAL:HG22	1:A:349:PHE:CD2	2.50	0.47
1:B:110:ARG:NH1	4:B:701:HOH:O	2.47	0.47
1:A:179:TYR:OH	1:A:209:HIS:NE2	2.30	0.47
1:A:160:ARG:HA	1:A:161:PRO:C	2.35	0.47
1:B:482:TYR:CE2	1:B:505:ARG:HB3	2.48	0.47
2:F:2:ASG:H6B	2:F:2:ASG:OSC	2.15	0.47
1:B:313:VAL:HG22	1:B:349:PHE:CD2	2.50	0.46
1:A:482:TYR:CE2	1:A:505:ARG:HB3	2.51	0.46
1:B:413:LEU:HD12	1:B:429:SER:HB3	1.96	0.46
1:A:149:PRO:HG3	1:A:166:TYR:CE2	2.51	0.46
1:B:27:ILE:HA	1:B:313:VAL:O	2.16	0.46
1:B:313:VAL:HG22	1:B:349:PHE:CG	2.50	0.46
1:C:315:ILE:HA	1:C:346:LEU:O	2.15	0.46
1:C:482:TYR:CE2	1:C:505:ARG:HB3	2.51	0.45
1:B:149:PRO:HG3	1:B:166:TYR:CE2	2.51	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:241:HIS:ND1	1:B:242:SER:O	2.43	0.45
1:C:27:ILE:HA	1:C:313:VAL:O	2.16	0.45
1:B:261:GLN:HG3	1:B:262:PRO:HD2	1.98	0.45
1:C:313:VAL:HG22	1:C:349:PHE:CG	2.52	0.45
1:A:313:VAL:HG22	1:A:349:PHE:CG	2.53	0.44
1:A:261:GLN:HG3	1:A:262:PRO:HD2	2.00	0.44
1:C:32:PRO:HD2	1:C:317:ALA:O	2.18	0.44
1:C:320:HIS:CE1	1:C:333:LYS:HE3	2.53	0.43
1:B:263:LEU:HD12	1:B:263:LEU:HA	1.83	0.43
1:A:32:PRO:HD2	1:A:317:ALA:O	2.19	0.43
1:A:263:LEU:HA	1:A:263:LEU:HD12	1.83	0.43
1:B:467:ASN:O	1:B:469:LEU:HD13	2.18	0.42
1:B:85:SER:N	1:B:86:PRO:CD	2.82	0.42
1:C:241:HIS:ND1	1:C:242:SER:O	2.42	0.42
1:A:278:LYS:O	1:A:281:SER:HB2	2.19	0.42
1:C:160:ARG:HA	1:C:161:PRO:C	2.39	0.42
1:C:448:THR:CG2	1:C:448:THR:O	2.68	0.42
2:F:1:BDP:O6B	2:F:1:BDP:O4	2.38	0.41
1:C:26:ASN:HB3	1:C:231:PHE:CE1	2.55	0.41
1:C:408:LEU:CD2	1:C:488:MET:HE3	2.50	0.41
1:C:445:ASP:OD1	1:C:448:THR:N	2.47	0.41
1:C:313:VAL:HG22	1:C:349:PHE:CD2	2.55	0.41
1:A:273:ASP:HB3	1:A:276:MET:HG3	2.03	0.41
1:A:393:LEU:HB2	1:A:401:ILE:HG21	2.03	0.41
1:A:74:LEU:HA	1:A:358:ASP:O	2.20	0.41
1:C:413:LEU:HD12	1:C:429:SER:HB3	2.02	0.41
1:A:205:TRP:CH2	1:A:291:GLY:HA3	2.55	0.41
1:C:136:PHE:HA	1:C:179:TYR:O	2.21	0.41
1:B:108:SER:HB3	1:B:162:VAL:HB	2.02	0.41
1:B:32:PRO:HD2	1:B:317:ALA:O	2.21	0.40
1:C:149:PRO:HG3	1:C:166:TYR:CE2	2.55	0.40
1:C:74:LEU:HA	1:C:358:ASP:O	2.20	0.40

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:60:HIS:NE2	1:C:60:HIS:NE2[1_455]	2.10	0.10

## 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	484/513 (94%)	451 (93%)	29 (6%)	4 (1%)	19	49
1	B	483/513 (94%)	449 (93%)	33 (7%)	1 (0%)	47	78
1	C	484/513 (94%)	448 (93%)	32 (7%)	4 (1%)	19	49
All	All	1451/1539 (94%)	1348 (93%)	94 (6%)	9 (1%)	25	56

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	159	GLN
1	B	159	GLN
1	C	184	PHE
1	C	159	GLN
1	C	500	LYS
1	C	503	SER
1	A	84	SER
1	A	184	PHE
1	A	503	SER

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

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	415/458 (91%)	391 (94%)	24 (6%)	20	50
1	B	413/458 (90%)	389 (94%)	24 (6%)	20	50
1	C	415/458 (91%)	395 (95%)	20 (5%)	25	58

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1243/1374 (90%)	1175 (94%)	68 (6%)	21 52

All (68) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	ILE
1	A	35	TYR
1	A	71	SER
1	A	99	ARG
1	A	104	LEU
1	A	109	THR
1	A	112	ILE
1	A	206	SER
1	A	223	ASN
1	A	244	TYR
1	A	261	GLN
1	A	263	LEU
1	A	280	GLU
1	A	333	LYS
1	A	384	GLU
1	A	388	ARG
1	A	401	ILE
1	A	437	ARG
1	A	454	SER
1	A	472	ASP
1	A	476	GLU
1	A	480	GLN
1	A	491	GLU
1	A	503	SER
1	B	27	ILE
1	B	28	ILE
1	B	35	TYR
1	B	71	SER
1	B	104	LEU
1	B	109	THR
1	B	112	ILE
1	B	160	ARG
1	B	206	SER
1	B	223	ASN
1	B	240	PRO
1	B	244	TYR
1	B	263	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	301	GLU
1	B	333	LYS
1	B	379	ASP
1	B	383	SER
1	B	388	ARG
1	B	401	ILE
1	B	448	THR
1	B	454	SER
1	B	472	ASP
1	B	480	GLN
1	B	491	GLU
1	C	27	ILE
1	C	46	GLU
1	C	55	ARG
1	C	71	SER
1	C	104	LEU
1	C	109	THR
1	C	112	ILE
1	C	201	ASP
1	C	206	SER
1	C	244	TYR
1	C	263	LEU
1	C	268	ILE
1	C	333	LYS
1	C	349	PHE
1	C	388	ARG
1	C	445	ASP
1	C	454	SER
1	C	472	ASP
1	C	480	GLN
1	C	503	SER

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

Mol	Chain	Res	Type
1	B	480	GLN
1	C	159	GLN

### 5.3.3 RNA

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](#)

9 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	BDP	D	1	2	10,13,13	1.35	1 (10%)	15,19,19	2.74	7 (46%)
2	ASG	D	2	3,2	18,18,19	1.23	1 (5%)	19,26,28	3.41	8 (42%)
2	GCD	D	3	2	7,11,12	3.11	4 (57%)	8,15,17	1.60	2 (25%)
2	BDP	E	1	2	10,13,13	2.09	3 (30%)	15,19,19	2.69	9 (60%)
2	ASG	E	2	3,2	18,18,19	1.05	2 (11%)	19,26,28	2.68	8 (42%)
2	GCD	E	3	2	7,11,12	2.22	1 (14%)	8,15,17	2.77	2 (25%)
2	BDP	F	1	2	10,13,13	0.88	0	15,19,19	2.55	7 (46%)
2	ASG	F	2	3,2	18,18,19	1.07	1 (5%)	19,26,28	2.91	8 (42%)
2	GCD	F	3	2	7,11,12	2.66	3 (42%)	8,15,17	1.75	3 (37%)

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	BDP	D	1	2	-	0/0/24/24	0/1/1/1
2	ASG	D	2	3,2	-	4/11/28/31	0/1/1/1
2	GCD	D	3	2	-	0/0/17/20	0/1/1/1
2	BDP	E	1	2	-	0/0/24/24	0/1/1/1
2	ASG	E	2	3,2	-	4/11/28/31	0/1/1/1
2	GCD	E	3	2	-	0/0/17/20	0/1/1/1
2	BDP	F	1	2	-	0/0/24/24	0/1/1/1
2	ASG	F	2	3,2	-	4/11/28/31	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GCD	F	3	2	-	0/0/17/20	0/1/1/1

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	3	GCD	O5-C5	5.69	1.45	1.37
2	D	3	GCD	O5-C5	5.48	1.45	1.37
2	F	3	GCD	O5-C5	5.33	1.44	1.37
2	D	3	GCD	O5-C1	-4.23	1.38	1.45
2	F	3	GCD	O5-C1	-3.66	1.39	1.45
2	E	1	BDP	C1-C2	3.60	1.60	1.52
2	D	3	GCD	C4-C5	3.27	1.36	1.32
2	E	1	BDP	O1-C1	3.00	1.49	1.39
2	E	1	BDP	O5-C5	2.70	1.47	1.44
2	D	1	BDP	O1-C1	2.70	1.48	1.39
2	D	3	GCD	C3-C4	2.61	1.53	1.50
2	F	3	GCD	O2-C2	2.46	1.48	1.43
2	F	2	ASG	OSA-S	2.27	1.64	1.50
2	E	2	ASG	OSA-S	2.20	1.63	1.50
2	E	2	ASG	O5-C1	2.10	1.47	1.43
2	D	2	ASG	OSA-S	2.02	1.62	1.50

All (54) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	2	ASG	OSA-S-OSC	-9.21	76.46	108.49
2	E	2	ASG	C4-O4-S	7.08	132.54	118.88
2	F	2	ASG	O5-C1-C2	-6.54	100.97	111.29
2	E	3	GCD	O5-C5-C4	-6.36	119.44	124.81
2	D	2	ASG	C8-C7-N2	5.98	126.22	116.10
2	F	2	ASG	C1-C2-N2	5.84	120.46	110.49
2	F	1	BDP	C6-C5-C4	-5.68	98.85	113.04
2	D	1	BDP	O5-C5-C4	5.38	117.76	108.91
2	D	2	ASG	C1-O5-C5	5.24	119.29	112.19
2	D	1	BDP	C3-C4-C5	5.05	119.38	109.02
2	F	2	ASG	C4-O4-S	4.96	128.46	118.88
2	E	2	ASG	O5-C1-C2	-4.63	103.98	111.29
2	D	2	ASG	OSC-S-OSB	-4.55	93.95	112.22
2	F	1	BDP	O4-C4-C5	-4.51	101.43	110.05
2	E	1	BDP	O1-C1-C2	4.18	120.81	109.03
2	E	1	BDP	O5-C5-C4	4.14	115.72	108.91
2	E	1	BDP	O2-C2-C3	-4.05	100.98	110.35

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	2	ASG	C1-C2-N2	4.01	117.34	110.49
2	E	1	BDP	C4-C3-C2	3.99	117.80	110.82
2	F	2	ASG	C8-C7-N2	3.97	122.83	116.10
2	F	1	BDP	O5-C1-C2	-3.97	103.21	110.28
2	D	2	ASG	O7-C7-C8	-3.93	114.77	122.06
2	E	3	GCD	O3-C3-C4	-3.90	100.55	109.31
2	F	2	ASG	O7-C7-C8	-3.74	115.12	122.06
2	D	1	BDP	O1-C1-C2	3.62	119.22	109.03
2	E	2	ASG	OSA-S-OSB	-3.51	96.28	108.49
2	D	1	BDP	C1-O5-C5	3.32	117.27	112.31
2	D	2	ASG	O5-C5-C6	3.24	112.28	107.20
2	F	2	ASG	C1-O5-C5	3.20	116.53	112.19
2	E	1	BDP	O2-C2-C1	3.13	116.41	109.16
2	E	2	ASG	O7-C7-N2	-3.06	116.33	121.95
2	D	2	ASG	OSA-S-OSB	3.04	119.07	108.49
2	D	3	GCD	O5-C5-C4	-3.02	122.26	124.81
2	D	1	BDP	O4-C4-C5	-2.91	104.48	110.05
2	D	1	BDP	O2-C2-C3	-2.86	103.75	110.35
2	E	1	BDP	C1-C2-C3	2.80	116.12	110.31
2	E	1	BDP	O3-C3-C2	-2.77	103.94	110.35
2	F	1	BDP	O3-C3-C4	-2.77	103.95	110.35
2	F	1	BDP	C1-C2-C3	-2.70	104.71	110.31
2	D	3	GCD	O2-C2-C1	2.68	114.63	109.15
2	E	2	ASG	C1-O5-C5	2.67	115.81	112.19
2	D	1	BDP	O3-C3-C2	2.67	116.51	110.35
2	E	1	BDP	C1-O5-C5	2.60	116.19	112.31
2	E	2	ASG	C8-C7-N2	2.56	120.43	116.10
2	D	2	ASG	C1-C2-N2	2.48	114.73	110.49
2	E	2	ASG	O5-C5-C4	2.41	114.58	110.07
2	F	3	GCD	C1-O5-C5	2.40	120.61	115.58
2	F	3	GCD	O2-C2-C1	2.37	114.00	109.15
2	F	2	ASG	OSA-S-OSB	-2.32	100.44	108.49
2	E	1	BDP	O4-C4-C5	2.26	114.37	110.05
2	F	1	BDP	O5-C5-C4	2.24	112.60	108.91
2	F	2	ASG	OSA-S-OSC	2.23	116.25	108.49
2	F	3	GCD	C3-C4-C5	-2.19	117.90	121.60
2	F	1	BDP	C3-C4-C5	-2.15	104.60	109.02

There are no chirality outliers.

All (12) torsion outliers are listed below:

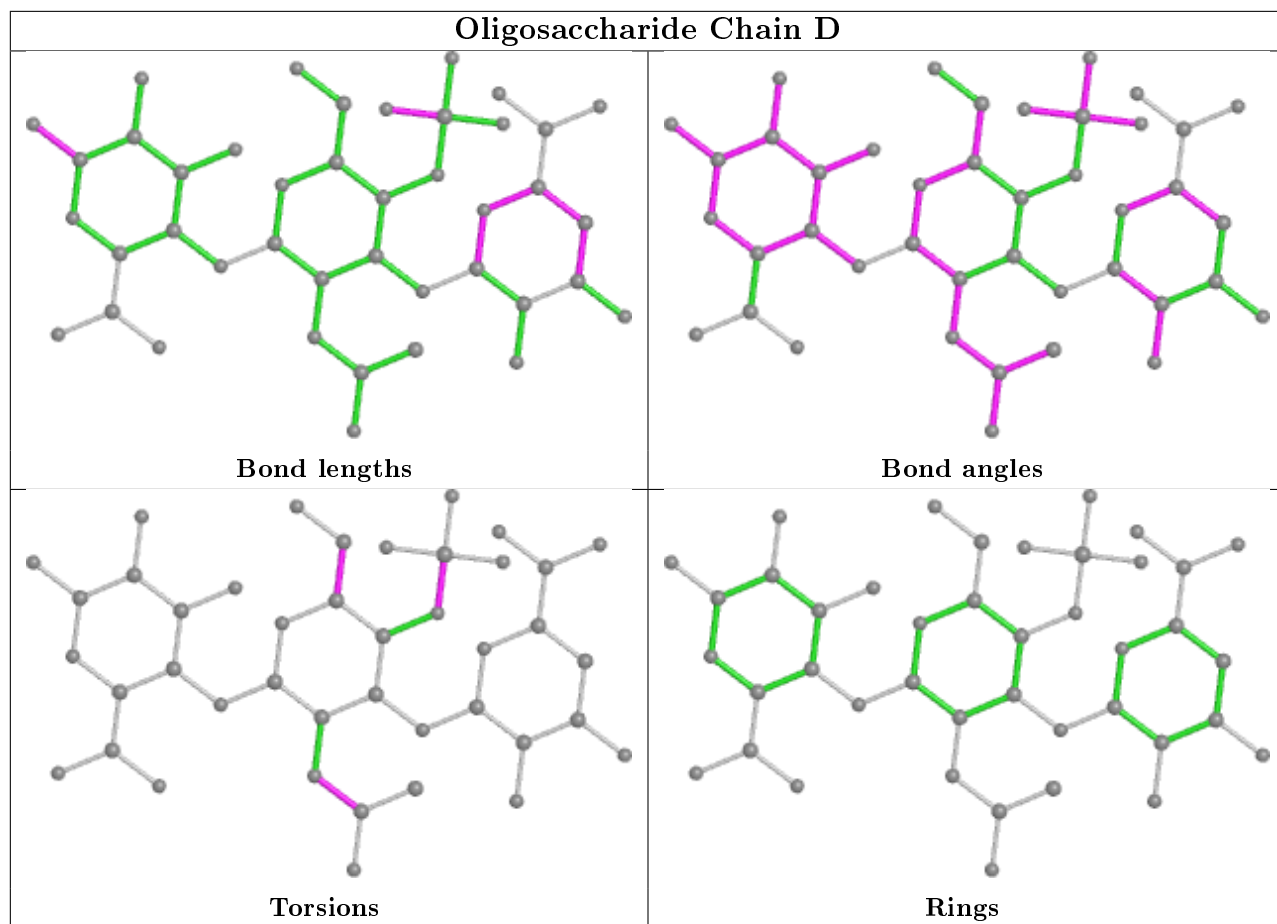
Mol	Chain	Res	Type	Atoms
2	E	2	ASG	C8-C7-N2-C2
2	E	2	ASG	O7-C7-N2-C2
2	D	2	ASG	C4-O4-S-OSA
2	F	2	ASG	C8-C7-N2-C2
2	F	2	ASG	O7-C7-N2-C2
2	E	2	ASG	C4-C5-C6-O6
2	D	2	ASG	C8-C7-N2-C2
2	D	2	ASG	O7-C7-N2-C2
2	F	2	ASG	C4-C5-C6-O6
2	F	2	ASG	O5-C5-C6-O6
2	E	2	ASG	O5-C5-C6-O6
2	D	2	ASG	O5-C5-C6-O6

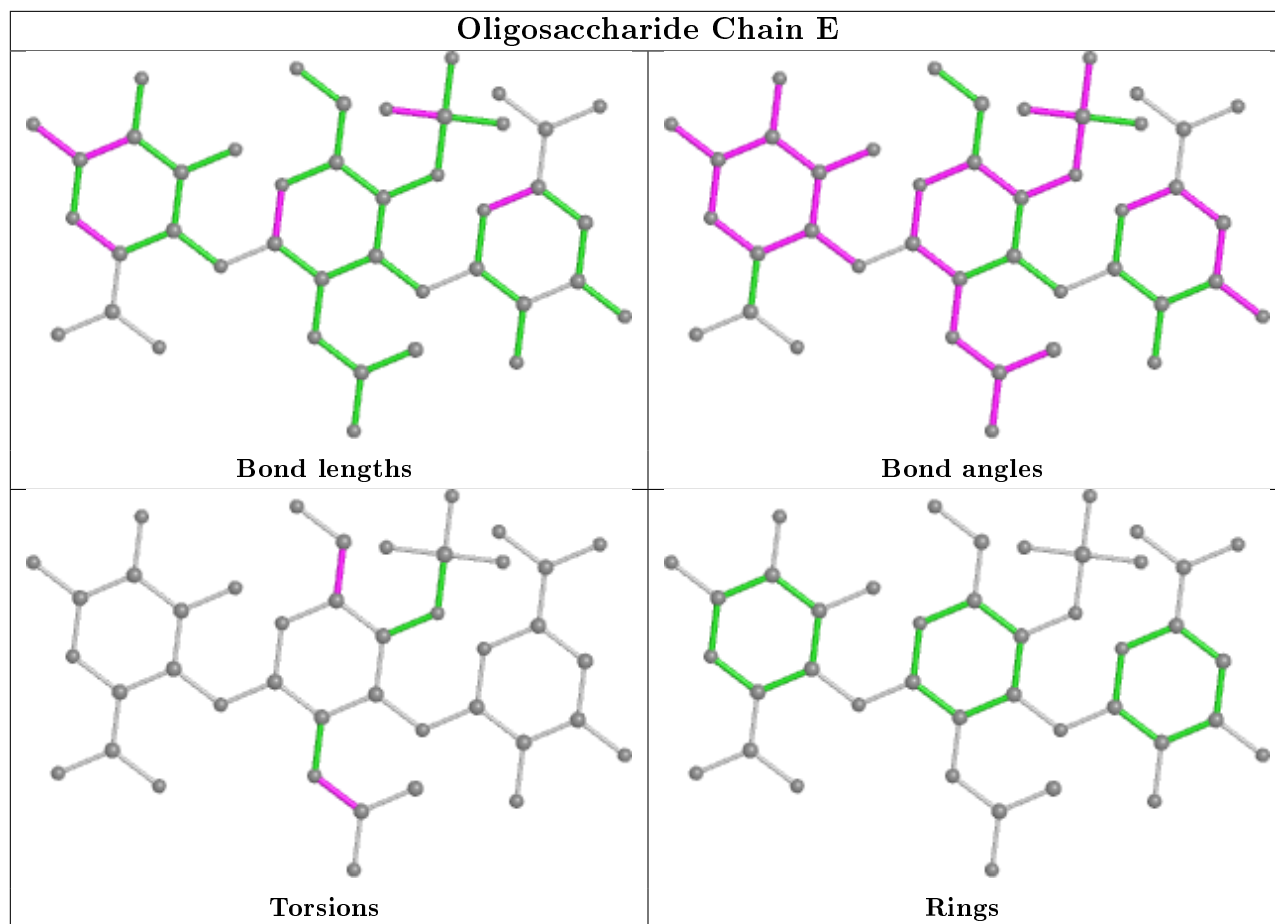
There are no ring outliers.

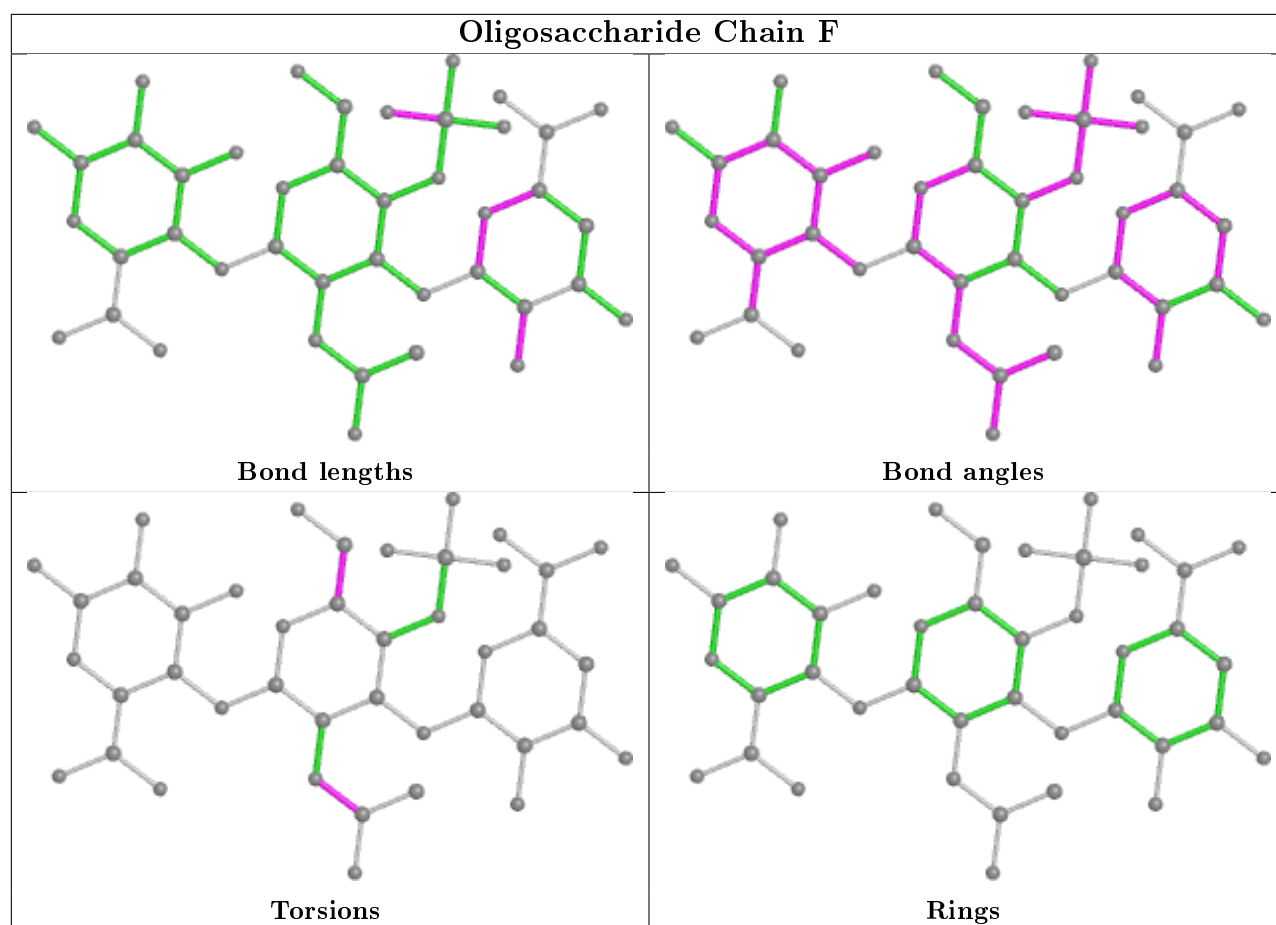
4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1	BDP	1	0
2	F	2	ASG	1	0
2	D	2	ASG	1	0
2	F	1	BDP	1	0

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







## 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

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.

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

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	486/513 (94%)	-0.53	0 100 100	20, 32, 50, 68	0
1	B	485/513 (94%)	-0.34	0 100 100	23, 38, 59, 72	0
1	C	486/513 (94%)	-0.44	0 100 100	22, 34, 49, 68	0
All	All	1457/1539 (94%)	-0.44	0 100 100	20, 35, 54, 72	0

There are no RSRZ outliers to report.

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

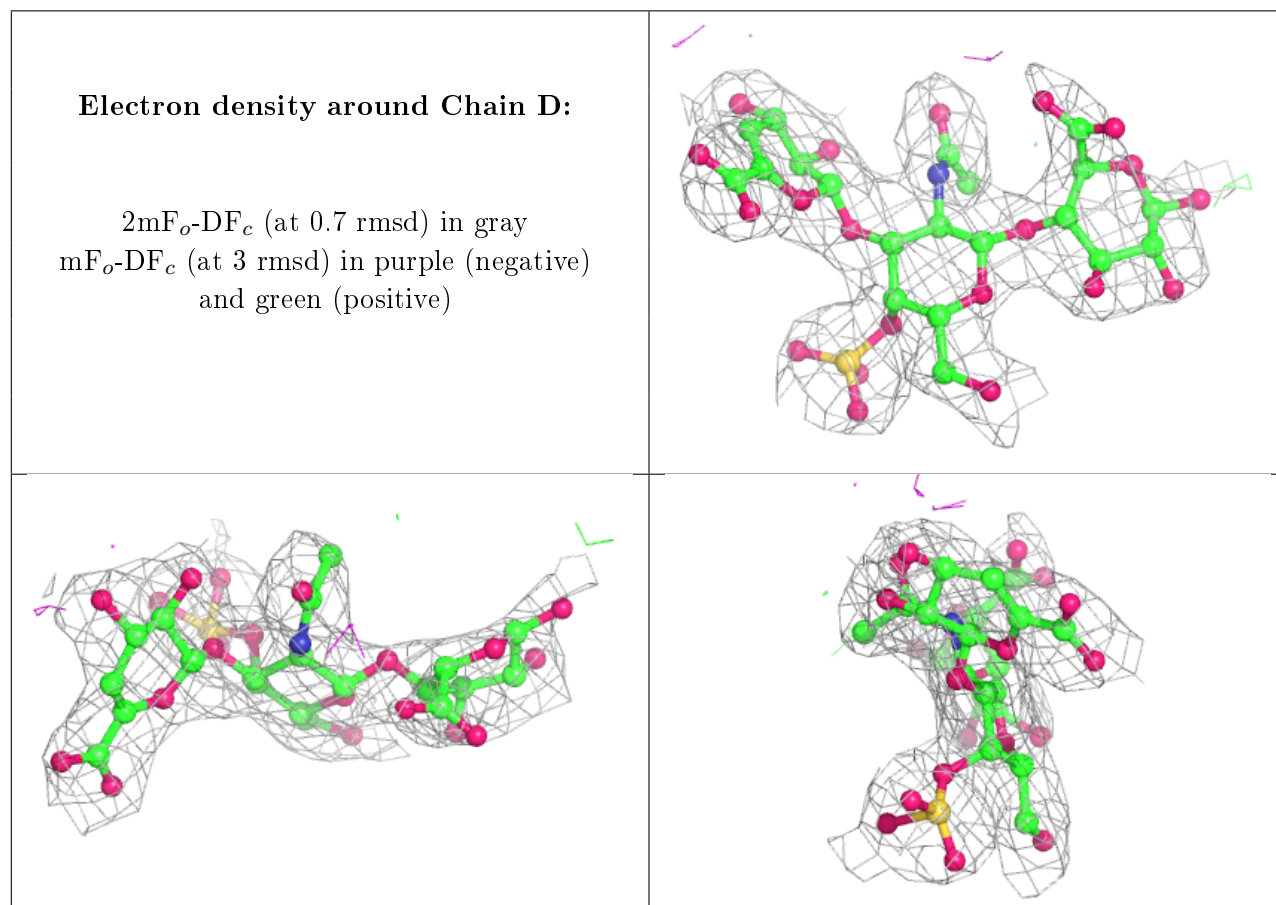
There are no non-standard protein/DNA/RNA residues in this entry.

### 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.

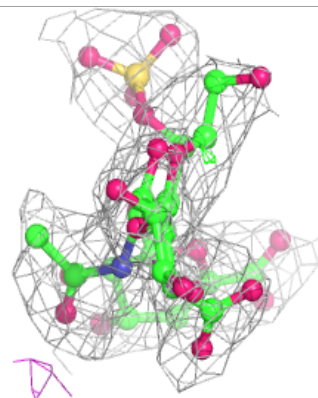
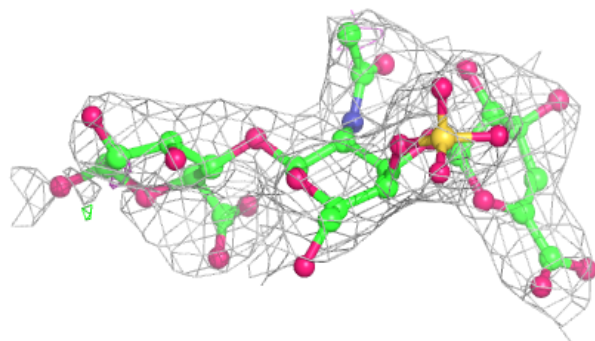
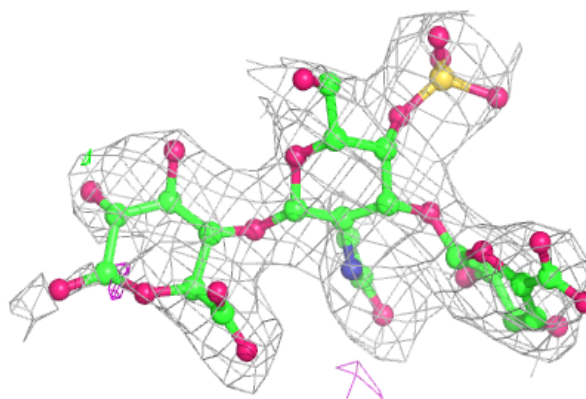
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	BDP	E	1	13/13	0.85	0.22	39,49,63,65	0
2	BDP	D	1	13/13	0.85	0.23	48,67,80,87	0
2	BDP	F	1	13/13	0.88	0.18	63,70,77,80	0
2	ASG	F	2	18/19	0.93	0.17	29,56,72,72	0
2	ASG	E	2	18/19	0.95	0.14	33,49,66,71	0
2	GCD	E	3	11/12	0.95	0.23	33,37,46,49	0
2	GCD	F	3	11/12	0.96	0.17	24,27,31,35	0
2	ASG	D	2	18/19	0.97	0.13	29,49,57,57	0
2	GCD	D	3	11/12	0.98	0.12	23,28,32,34	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.

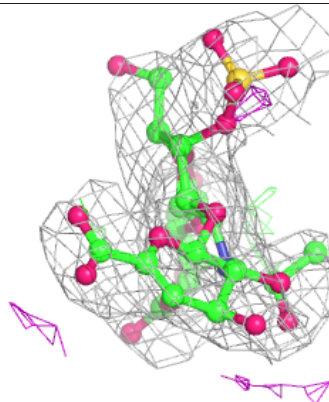
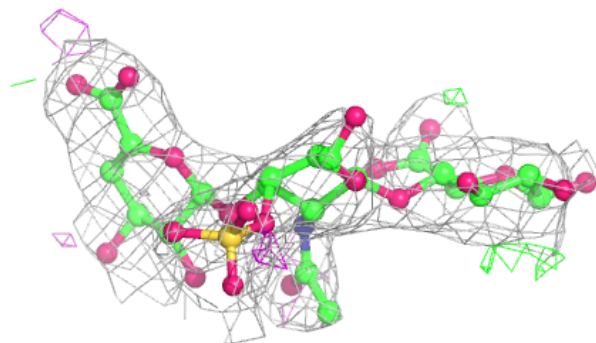
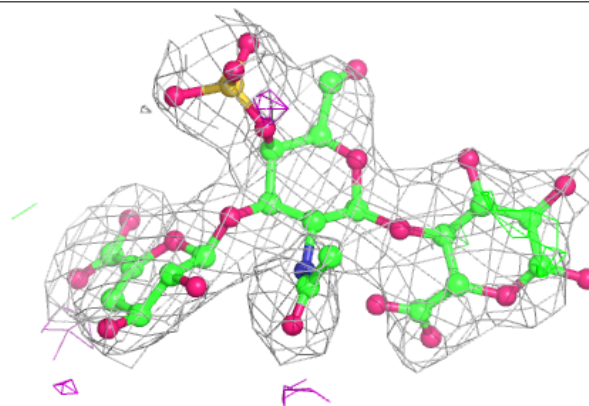


**Electron density around Chain E:**

$2mF_o-DF_c$  (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)



## 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
3	CA	A	601	1/1	0.88	0.08	47,47,47,47	0
3	CA	B	601	1/1	0.97	0.13	63,63,63,63	0
3	CA	C	601	1/1	0.97	0.06	52,52,52,52	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.