



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 8, 2020 – 01:17 PM BST

PDB ID : 4H04  
Title : Lacto-N-biosidase from Bifidobacterium bifidum  
Authors : Ito, T.; Katayama, T.; Wada, J.; Suzuki, R.; Ashida, H.; Wakagi, T.; Yamamoto, K.; Fushinobu, S.  
Deposited on : 2012-09-07  
Resolution : 1.80 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

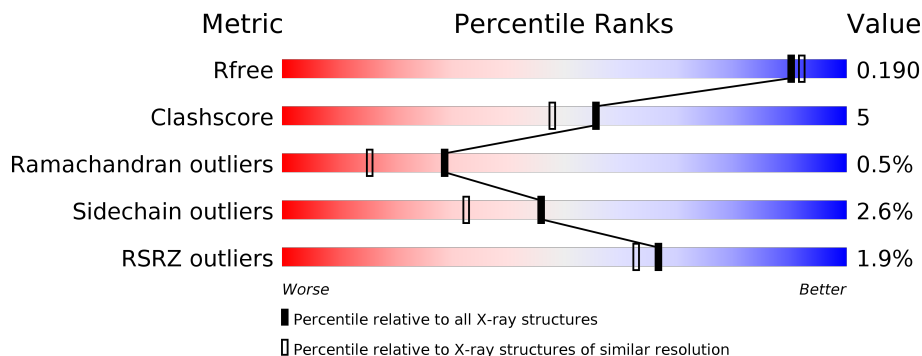
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	644	<p>2% 82% 14%</p>
1	B	644	<p>2% 84% 13%</p>
2	C	2	<p>50% 50%</p>
2	D	2	<p>100%</p>

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 11351 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 Lacto-N-biosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	633	4969	3132	846	974	17	0	0	0
1	B	632	4963	3129	845	972	17	0	0	0

There are 42 discrepancies between the modelled and reference sequences:

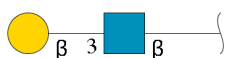
Chain	Residue	Modelled	Actual	Comment	Reference
A	20	MET	-	expression tag	UNP B3TLD6
A	21	GLY	-	expression tag	UNP B3TLD6
A	22	SER	-	expression tag	UNP B3TLD6
A	23	SER	-	expression tag	UNP B3TLD6
A	24	HIS	-	expression tag	UNP B3TLD6
A	25	HIS	-	expression tag	UNP B3TLD6
A	26	HIS	-	expression tag	UNP B3TLD6
A	27	HIS	-	expression tag	UNP B3TLD6
A	28	HIS	-	expression tag	UNP B3TLD6
A	29	HIS	-	expression tag	UNP B3TLD6
A	30	SER	-	expression tag	UNP B3TLD6
A	31	SER	-	expression tag	UNP B3TLD6
A	32	GLY	-	expression tag	UNP B3TLD6
A	33	LEU	-	expression tag	UNP B3TLD6
A	34	VAL	-	expression tag	UNP B3TLD6
A	35	PRO	-	expression tag	UNP B3TLD6
A	36	ARG	-	expression tag	UNP B3TLD6
A	37	GLY	-	expression tag	UNP B3TLD6
A	38	SER	-	expression tag	UNP B3TLD6
A	39	HIS	-	expression tag	UNP B3TLD6
A	40	MET	-	expression tag	UNP B3TLD6
B	20	MET	-	expression tag	UNP B3TLD6
B	21	GLY	-	expression tag	UNP B3TLD6
B	22	SER	-	expression tag	UNP B3TLD6
B	23	SER	-	expression tag	UNP B3TLD6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	24	HIS	-	expression tag	UNP B3TLD6
B	25	HIS	-	expression tag	UNP B3TLD6
B	26	HIS	-	expression tag	UNP B3TLD6
B	27	HIS	-	expression tag	UNP B3TLD6
B	28	HIS	-	expression tag	UNP B3TLD6
B	29	HIS	-	expression tag	UNP B3TLD6
B	30	SER	-	expression tag	UNP B3TLD6
B	31	SER	-	expression tag	UNP B3TLD6
B	32	GLY	-	expression tag	UNP B3TLD6
B	33	LEU	-	expression tag	UNP B3TLD6
B	34	VAL	-	expression tag	UNP B3TLD6
B	35	PRO	-	expression tag	UNP B3TLD6
B	36	ARG	-	expression tag	UNP B3TLD6
B	37	GLY	-	expression tag	UNP B3TLD6
B	38	SER	-	expression tag	UNP B3TLD6
B	39	HIS	-	expression tag	UNP B3TLD6
B	40	MET	-	expression tag	UNP B3TLD6

- Molecule 2 is an oligosaccharide called beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	2	26	14	1	11	0	0	0
2	D	2	26	14	1	11	0	0	0

- 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	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		

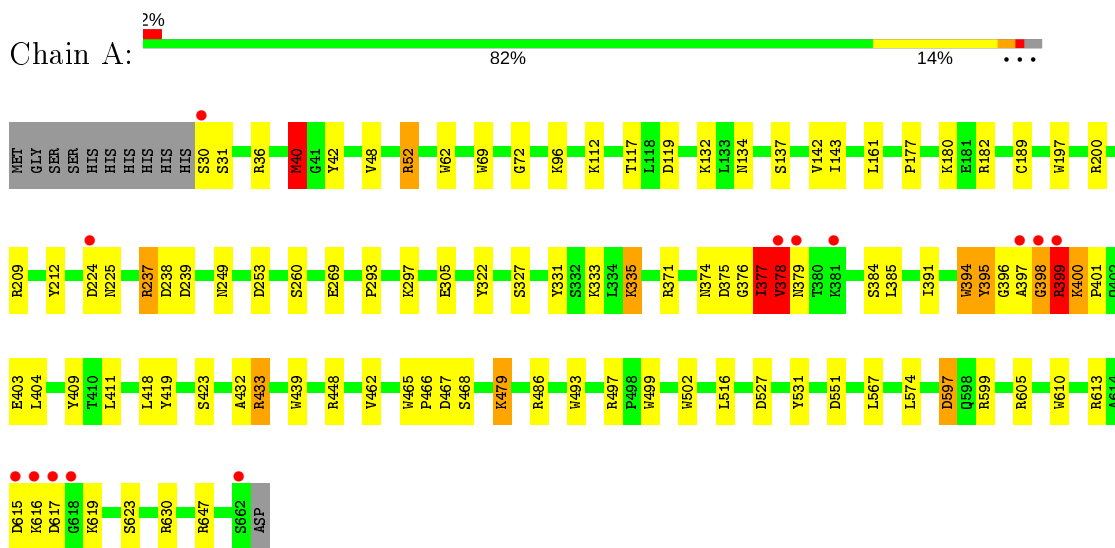
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	656	Total	O	0	0
			656	656		
4	B	671	Total	O	0	0
			671	671		

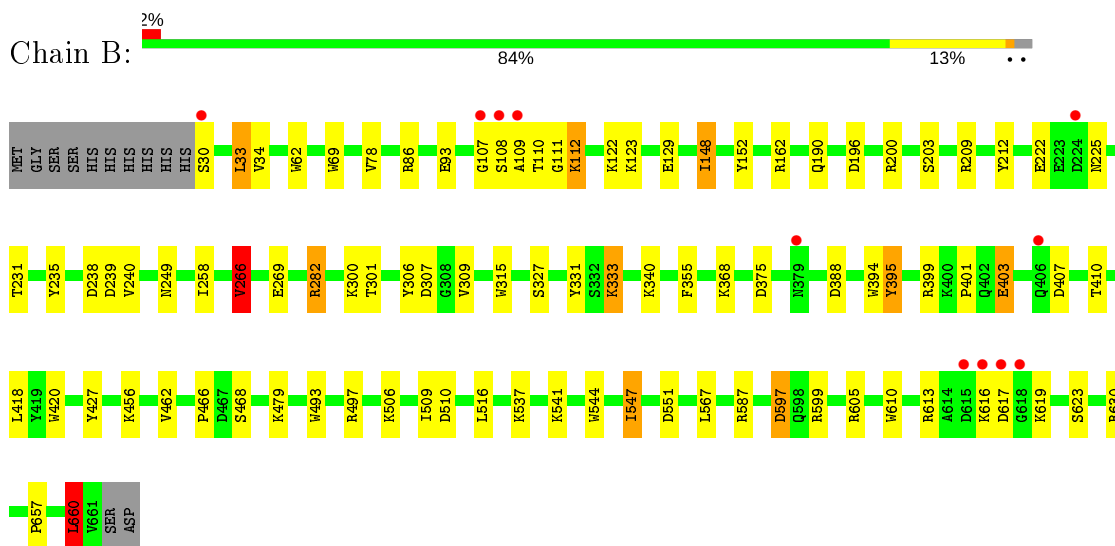
### 3 Residue-property plots [i](#)

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: Lacto-N-biosidase



- Molecule 1: Lacto-N-biosidase



- Molecule 2: beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose



HAG1  
GAL2

- Molecule 2: beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:

100%

HAG1  
GAL2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	116.81Å 131.03Å 104.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.51 – 1.80 47.51 – 1.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (47.51-1.80) 99.9 (47.51-1.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.69 (at 1.79Å)	Xtrriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.155 , 0.191 0.154 , 0.190	Depositor DCC
$R_{free}$ test set	7448 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.9	Xtrriage
Anisotropy	0.023	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 54.2	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	0.96	EDS
Total number of atoms	11351	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.51% 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: GAL, NAG, SO4

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	1.41	23/5078 (0.5%)	1.36	38/6891 (0.6%)
1	B	1.42	20/5072 (0.4%)	1.31	33/6883 (0.5%)
All	All	1.42	43/10150 (0.4%)	1.34	71/13774 (0.5%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

All (43) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	497	ARG	CZ-NH2	9.92	1.46	1.33
1	A	62	TRP	CD2-CE2	9.03	1.52	1.41
1	B	62	TRP	CD2-CE2	8.65	1.51	1.41
1	A	394	TRP	CD2-CE2	7.58	1.50	1.41
1	A	409	TYR	CE1-CZ	7.36	1.48	1.38
1	A	62	TRP	CG-CD1	7.04	1.46	1.36
1	A	52	ARG	CZ-NH1	6.90	1.42	1.33
1	B	493	TRP	CD2-CE2	6.76	1.49	1.41
1	A	465	TRP	CD2-CE2	6.60	1.49	1.41
1	B	307	ASP	CB-CG	-6.57	1.38	1.51
1	B	420	TRP	CD2-CE2	6.50	1.49	1.41
1	B	269	GLU	CG-CD	6.49	1.61	1.51
1	B	162	ARG	CZ-NH1	6.44	1.41	1.33
1	B	315	TRP	CD2-CE2	6.30	1.49	1.41
1	A	597	ASP	CB-CG	6.27	1.65	1.51
1	A	419	TYR	CG-CD1	5.99	1.47	1.39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	260	SER	CA-CB	5.93	1.61	1.52
1	A	69	TRP	CD2-CE2	5.91	1.48	1.41
1	A	439	TRP	CD2-CE2	5.82	1.48	1.41
1	A	395	TYR	CB-CG	5.80	1.60	1.51
1	A	189	CYS	CB-SG	5.80	1.92	1.82
1	A	395	TYR	CE2-CZ	-5.77	1.31	1.38
1	A	269	GLU	CD-OE2	-5.76	1.19	1.25
1	A	371	ARG	CZ-NH1	5.74	1.40	1.33
1	B	203	SER	CB-OG	-5.54	1.35	1.42
1	A	610	TRP	CD2-CE2	5.51	1.48	1.41
1	B	427	TYR	CG-CD1	5.43	1.46	1.39
1	B	610	TRP	CD2-CE2	5.43	1.47	1.41
1	B	544	TRP	CD2-CE2	5.35	1.47	1.41
1	B	605	ARG	CZ-NH2	5.33	1.40	1.33
1	B	599	ARG	CZ-NH2	5.32	1.40	1.33
1	B	605	ARG	CZ-NH1	5.30	1.40	1.33
1	A	197	TRP	CE3-CZ3	5.25	1.47	1.38
1	A	499	TRP	CE3-CZ3	5.25	1.47	1.38
1	B	69	TRP	CG-CD1	5.23	1.44	1.36
1	A	269	GLU	CB-CG	-5.21	1.42	1.52
1	A	384	SER	CB-OG	-5.20	1.35	1.42
1	B	269	GLU	CD-OE1	-5.19	1.20	1.25
1	B	395	TYR	CB-CG	5.14	1.59	1.51
1	A	331	TYR	CE1-CZ	5.12	1.45	1.38
1	B	597	ASP	CB-CG	5.11	1.62	1.51
1	A	31	SER	C-O	5.11	1.33	1.23
1	B	331	TYR	CE1-CZ	5.03	1.45	1.38

All (71) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	237	ARG	NE-CZ-NH2	-19.67	110.46	120.30
1	A	52	ARG	NE-CZ-NH1	17.60	129.10	120.30
1	A	433	ARG	NE-CZ-NH1	15.52	128.06	120.30
1	A	237	ARG	NE-CZ-NH1	15.09	127.84	120.30
1	A	52	ARG	NE-CZ-NH2	-13.91	113.35	120.30
1	B	605	ARG	NE-CZ-NH2	-12.92	113.84	120.30
1	B	239	ASP	CB-CG-OD2	-11.91	107.58	118.30
1	B	266	VAL	CG1-CB-CG2	11.67	129.58	110.90
1	B	497	ARG	NE-CZ-NH1	-10.36	115.12	120.30
1	B	239	ASP	CB-CG-OD1	9.87	127.18	118.30
1	B	200	ARG	NE-CZ-NH1	9.51	125.05	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	239	ASP	CB-CG-OD2	-9.21	110.01	118.30
1	A	239	ASP	CB-CG-OD1	9.18	126.56	118.30
1	A	433	ARG	NE-CZ-NH2	-8.68	115.96	120.30
1	A	486	ARG	NE-CZ-NH1	8.33	124.47	120.30
1	B	86	ARG	NE-CZ-NH1	-8.21	116.19	120.30
1	A	237	ARG	CG-CD-NE	-8.17	94.64	111.80
1	A	377	ILE	CG1-CB-CG2	-7.95	93.92	111.40
1	B	660	LEU	CB-CG-CD1	7.79	124.25	111.00
1	A	551	ASP	CB-CG-OD1	7.66	125.19	118.30
1	B	613	ARG	NE-CZ-NH2	-7.37	116.62	120.30
1	B	605	ARG	NE-CZ-NH1	7.21	123.91	120.30
1	A	40	MET	CG-SD-CE	7.08	111.53	100.20
1	A	479	LYS	CD-CE-NZ	-6.57	96.59	111.70
1	A	322	TYR	CB-CG-CD2	-6.55	117.07	121.00
1	A	404	LEU	CB-CG-CD2	-6.54	99.88	111.00
1	A	433	ARG	CD-NE-CZ	6.49	132.69	123.60
1	A	599	ARG	NE-CZ-NH2	-6.38	117.11	120.30
1	B	152	TYR	CD1-CE1-CZ	-6.38	114.06	119.80
1	B	516	LEU	CB-CG-CD2	-6.18	100.49	111.00
1	A	297	LYS	CD-CE-NZ	-6.16	97.54	111.70
1	A	200	ARG	NE-CZ-NH1	6.09	123.34	120.30
1	A	305	GLU	OE1-CD-OE2	-6.06	116.03	123.30
1	A	605	ARG	NE-CZ-NH2	-5.95	117.33	120.30
1	B	300	LYS	CD-CE-NZ	-5.95	98.01	111.70
1	A	409	TYR	CB-CG-CD2	-5.93	117.44	121.00
1	B	129	GLU	OE1-CD-OE2	5.93	130.42	123.30
1	B	375	ASP	N-CA-C	5.93	127.02	111.00
1	B	599	ARG	NE-CZ-NH2	-5.85	117.38	120.30
1	B	355	PHE	CB-CG-CD2	-5.82	116.73	120.80
1	A	613	ARG	NE-CZ-NH2	-5.75	117.42	120.30
1	A	333	LYS	CD-CE-NZ	5.75	124.92	111.70
1	A	182	ARG	NE-CZ-NH1	-5.74	117.43	120.30
1	A	375	ASP	N-CA-C	5.74	126.49	111.00
1	B	479	LYS	CD-CE-NZ	-5.74	98.51	111.70
1	B	307	ASP	CB-CG-OD2	-5.72	113.15	118.30
1	B	162	ARG	NE-CZ-NH1	5.71	123.15	120.30
1	A	597	ASP	CB-CG-OD1	5.65	123.39	118.30
1	A	497	ARG	NE-CZ-NH2	5.65	123.12	120.30
1	B	551	ASP	CB-CG-OD1	5.55	123.30	118.30
1	B	148	ILE	CG1-CB-CG2	-5.54	99.20	111.40
1	A	419	TYR	CB-CG-CD1	5.51	124.31	121.00
1	B	152	TYR	CB-CG-CD2	-5.51	117.69	121.00

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	388	ASP	CB-CG-OD2	-5.50	113.35	118.30
1	B	196	ASP	CB-CG-OD2	-5.49	113.36	118.30
1	A	527	ASP	CB-CG-OD1	5.48	123.23	118.30
1	B	238	ASP	CB-CG-OD1	5.43	123.19	118.30
1	B	547	ILE	CB-CG1-CD1	-5.40	98.78	113.90
1	A	419	TYR	CB-CG-CD2	-5.38	117.77	121.00
1	B	510	ASP	CB-CG-OD1	5.35	123.12	118.30
1	B	231	THR	CA-CB-CG2	-5.34	104.92	112.40
1	A	647	ARG	NE-CZ-NH1	5.28	122.94	120.30
1	A	161	LEU	CA-CB-CG	-5.27	103.17	115.30
1	B	613	ARG	NE-CZ-NH1	5.27	122.94	120.30
1	B	112	LYS	CD-CE-NZ	5.25	123.76	111.70
1	A	399	ARG	N-CA-C	5.24	125.15	111.00
1	B	235	TYR	CD1-CE1-CZ	-5.21	115.11	119.80
1	B	587	ARG	NE-CZ-NH1	5.20	122.90	120.30
1	A	253	ASP	CB-CG-OD2	-5.17	113.65	118.30
1	A	224	ASP	CB-CG-OD2	-5.10	113.71	118.30
1	A	119	ASP	CB-CG-OD2	-5.04	113.76	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	468	SER	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	4969	0	4845	53	0
1	B	4963	0	4840	47	0
2	C	26	0	24	2	0
2	D	26	0	24	0	0
3	A	20	0	0	0	0
3	B	20	0	0	0	0
4	A	656	0	0	11	0
4	B	671	0	0	14	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	11351	0	9733	100	0

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

All (100) 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:399:ARG:HB3	1:A:403:GLU:OE1	1.48	1.14
1:A:378:VAL:HG23	4:A:1007:HOH:O	1.52	1.07
1:B:597:ASP:HB3	4:B:1387:HOH:O	1.58	1.04
1:B:109:ALA:HB3	4:B:1417:HOH:O	1.61	0.98
1:A:180:LYS:HD3	4:A:915:HOH:O	1.64	0.95
1:B:282:ARG:HG3	1:B:282:ARG:HH11	1.30	0.92
1:A:400:LYS:HD2	1:A:448:ARG:HE	1.37	0.88
1:B:33:LEU:HD23	1:B:33:LEU:C	1.93	0.88
1:A:597:ASP:HB3	4:A:1086:HOH:O	1.73	0.88
1:A:376:GLY:O	1:A:378:VAL:HG22	1.76	0.84
1:B:282:ARG:CG	1:B:282:ARG:HH11	1.92	0.82
1:A:237:ARG:HD3	4:A:1277:HOH:O	1.79	0.82
1:B:282:ARG:NH1	1:B:282:ARG:HG3	1.96	0.78
1:B:333:LYS:HD2	4:B:1084:HOH:O	1.83	0.77
1:B:547:ILE:HG12	4:B:1007:HOH:O	1.85	0.77
1:A:40:MET:HG2	1:A:112:LYS:HG2	1.68	0.76
1:B:222:GLU:HG2	4:B:1200:HOH:O	1.84	0.76
1:B:78:VAL:CG1	1:B:107:GLY:HA2	2.16	0.76
1:A:52:ARG:NH2	4:A:1241:HOH:O	2.22	0.73
1:A:48:VAL:HG23	4:A:902:HOH:O	1.90	0.70
1:A:400:LYS:CD	1:A:448:ARG:HE	2.05	0.69
1:A:117:THR:HB	1:A:143:ILE:HD12	1.75	0.69
1:B:110:THR:HG22	1:B:111:GLY:N	2.09	0.68
1:A:615:ASP:O	1:A:617:ASP:HA	1.94	0.67
1:A:238:ASP:HB3	4:A:1432:HOH:O	1.95	0.65
1:B:340:LYS:HE2	4:B:1423:HOH:O	1.96	0.65
1:B:78:VAL:HG13	1:B:107:GLY:HA2	1.80	0.64
1:A:376:GLY:O	1:A:378:VAL:CG2	2.46	0.64
1:B:399:ARG:HH21	1:B:403:GLU:CD	2.01	0.64
1:A:40:MET:SD	1:A:112:LYS:HE2	2.39	0.62
1:A:374:ASN:O	1:A:377:ILE:HG22	2.00	0.61
1:A:401:PRO:CD	1:A:448:ARG:HD3	2.30	0.61
1:A:293:PRO:HG2	4:A:1168:HOH:O	2.01	0.61

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:33:LEU:C	1:B:33:LEU:CD2	2.69	0.59
1:A:401:PRO:HD2	1:A:448:ARG:HD3	1.84	0.58
1:B:225:ASN:HD21	1:B:301:THR:HG21	1.67	0.58
1:A:615:ASP:OD1	1:A:619:LYS:HB2	2.04	0.58
1:A:400:LYS:HD2	1:A:448:ARG:NE	2.13	0.57
1:B:333:LYS:HE3	4:B:1091:HOH:O	2.06	0.56
1:B:368:LYS:HD3	4:B:1350:HOH:O	2.06	0.56
1:B:78:VAL:HG11	1:B:109:ALA:HB2	1.90	0.54
1:B:108:SER:HA	4:B:1412:HOH:O	2.09	0.53
1:B:240:VAL:CG1	1:B:309:VAL:CG1	2.88	0.51
1:A:423:SER:HB3	1:A:468:SER:HB3	1.94	0.50
2:C:1:NAG:H1	2:C:1:NAG:O7	2.12	0.50
1:A:377:ILE:HG13	1:A:385:LEU:HD22	1.94	0.49
1:B:78:VAL:HG13	1:B:107:GLY:CA	2.43	0.49
1:B:78:VAL:HG11	1:B:107:GLY:HA2	1.92	0.48
1:A:132:LYS:HE2	1:A:134:ASN:OD1	2.13	0.48
1:B:110:THR:CG2	1:B:111:GLY:N	2.76	0.48
1:B:340:LYS:CE	4:B:1423:HOH:O	2.60	0.48
1:A:40:MET:HG2	1:A:112:LYS:CG	2.42	0.47
1:A:391:ILE:HB	1:A:411:LEU:HD23	1.96	0.47
1:B:537:LYS:NZ	1:B:541:LYS:HD2	2.29	0.47
1:A:249:ASN:HB3	4:A:1224:HOH:O	2.14	0.47
1:B:657:PRO:HD2	1:B:660:LEU:HD22	1.96	0.47
1:B:78:VAL:CG1	1:B:107:GLY:CA	2.91	0.46
1:A:96:LYS:HD2	1:A:516:LEU:HD21	1.98	0.46
1:B:110:THR:HG22	1:B:111:GLY:H	1.75	0.46
1:B:258:ILE:HD11	1:B:306:TYR:CE1	2.50	0.46
1:A:40:MET:HB3	1:A:72:GLY:HA3	1.97	0.46
1:B:623:SER:OG	1:B:630:ARG:HD3	2.15	0.46
1:A:143:ILE:HD11	4:A:1134:HOH:O	2.14	0.46
1:B:33:LEU:HD23	1:B:34:VAL:N	2.30	0.46
1:A:418:LEU:HB3	1:A:462:VAL:HA	1.97	0.45
1:B:249:ASN:HB3	4:B:1040:HOH:O	2.16	0.45
1:A:177:PRO:HD3	1:A:493:TRP:CH2	2.52	0.45
1:B:597:ASP:CB	4:B:1387:HOH:O	2.39	0.45
1:A:398:GLY:HA2	1:A:399:ARG:HA	1.31	0.45
1:A:42:TYR:CG	1:A:137:SER:HA	2.52	0.45
1:B:190:GLN:HG2	1:B:266:VAL:HG21	1.99	0.45
1:A:623:SER:OG	1:A:630:ARG:HD3	2.17	0.45
1:A:142:VAL:C	1:A:143:ILE:HD13	2.37	0.44
1:A:40:MET:CG	1:A:112:LYS:HE2	2.47	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:395:TYR:O	1:A:397:ALA:N	2.50	0.44
1:B:418:LEU:HB3	1:B:462:VAL:HA	1.98	0.44
1:A:432:ALA:HA	1:A:502:TRP:CG	2.53	0.44
1:A:531:TYR:CE1	1:B:33:LEU:HD21	2.53	0.44
1:B:506:LYS:HA	1:B:509:ILE:HG12	2.00	0.44
1:A:377:ILE:O	1:A:377:ILE:HG23	2.17	0.43
1:A:394:TRP:CD2	1:A:395:TYR:HB2	2.53	0.43
1:B:394:TRP:CD2	1:B:395:TYR:HB2	2.53	0.43
1:A:335:LYS:HG3	1:A:335:LYS:HZ2	1.56	0.43
1:A:399:ARG:HB2	1:A:403:GLU:HB2	2.01	0.42
1:B:399:ARG:NH2	1:B:407:ASP:OD2	2.44	0.42
1:A:467:ASP:O	1:A:574:LEU:HB3	2.19	0.42
2:C:1:NAG:O7	2:C:1:NAG:C1	2.68	0.42
1:A:423:SER:CB	1:A:468:SER:HB3	2.49	0.42
1:A:479:LYS:HE3	4:A:1014:HOH:O	2.21	0.41
1:A:378:VAL:HB	1:A:379:ASN:H	1.16	0.41
1:B:410:THR:HA	1:B:456:LYS:HB3	2.02	0.41
1:B:148:ILE:HD12	1:B:148:ILE:HG23	1.86	0.41
1:B:401:PRO:HD2	4:B:1152:HOH:O	2.20	0.41
1:A:40:MET:HG2	1:A:112:LYS:HE2	2.03	0.41
1:B:122:LYS:HE2	4:B:1018:HOH:O	2.20	0.41
1:A:212:TYR:CD2	1:A:212:TYR:C	2.94	0.41
1:A:30:SER:OG	1:B:619:LYS:HD3	2.21	0.41
1:B:225:ASN:N	1:B:225:ASN:OD1	2.35	0.40
1:A:615:ASP:HB2	1:A:619:LYS:H	1.86	0.40
1:B:240:VAL:HG11	1:B:309:VAL:CG1	2.51	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	631/644 (98%)	611 (97%)	16 (2%)	4 (1%)	25 12
1	B	630/644 (98%)	608 (96%)	20 (3%)	2 (0%)	41 27
All	All	1261/1288 (98%)	1219 (97%)	36 (3%)	6 (0%)	29 15

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	398	GLY
1	A	396	GLY
1	B	617	ASP
1	A	466	PRO
1	A	378	VAL
1	B	466	PRO

### 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	530/540 (98%)	517 (98%)	13 (2%)	47 34
1	B	529/540 (98%)	514 (97%)	15 (3%)	43 30
All	All	1059/1080 (98%)	1031 (97%)	28 (3%)	46 32

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	36	ARG
1	A	40	MET
1	A	209	ARG
1	A	225	ASN
1	A	327	SER
1	A	335	LYS
1	A	377	ILE
1	A	378	VAL
1	A	399	ARG
1	A	400	LYS

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	433	ARG
1	A	567	LEU
1	A	616	LYS
1	B	30	SER
1	B	33	LEU
1	B	93	GLU
1	B	112	LYS
1	B	123	LYS
1	B	209	ARG
1	B	212	TYR
1	B	266	VAL
1	B	282	ARG
1	B	327	SER
1	B	333	LYS
1	B	403	GLU
1	B	567	LEU
1	B	616	LYS
1	B	660	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

4 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	NAG	C	1	2	15,15,15	1.31	2 (13%)	21,21,21	3.39	9 (42%)
2	GAL	C	2	2	11,11,12	0.79	0	15,15,17	1.75	3 (20%)
2	NAG	D	1	2	15,15,15	1.29	2 (13%)	21,21,21	2.45	4 (19%)
2	GAL	D	2	2	11,11,12	1.25	2 (18%)	15,15,17	1.44	3 (20%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	C	1	2	-	2/6/26/26	0/1/1/1
2	GAL	C	2	2	-	0/2/19/22	0/1/1/1
2	NAG	D	1	2	-	2/6/26/26	0/1/1/1
2	GAL	D	2	2	-	0/2/19/22	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	NAG	C2-N2	2.77	1.50	1.45
2	D	1	NAG	C8-C7	2.75	1.56	1.50
2	C	1	NAG	O1-C1	2.69	1.48	1.39
2	D	2	GAL	C4-C3	2.25	1.58	1.52
2	D	2	GAL	C6-C5	2.08	1.58	1.51
2	D	1	NAG	O1-C1	2.02	1.46	1.39

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1	NAG	O5-C1-C2	11.53	121.11	109.52
2	D	1	NAG	C1-C2-C3	6.75	119.75	110.54
2	D	1	NAG	C1-O5-C5	5.55	124.13	113.66
2	C	1	NAG	C1-C2-C3	4.84	117.14	110.54
2	C	2	GAL	O5-C5-C6	4.54	114.31	107.20
2	C	1	NAG	C1-C2-N2	3.80	115.13	110.73
2	D	1	NAG	O5-C1-C2	3.61	113.14	109.52
2	C	1	NAG	C2-N2-C7	-3.46	114.75	123.18
2	C	2	GAL	O2-C2-C3	-3.40	103.33	110.14
2	C	1	NAG	O5-C5-C4	3.22	115.53	109.69
2	C	1	NAG	C8-C7-N2	-3.15	110.77	116.10
2	C	1	NAG	C3-C4-C5	-3.13	104.65	110.24

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1	NAG	C3-C4-C5	-2.89	105.09	110.24
2	D	2	GAL	O2-C2-C3	-2.86	104.42	110.14
2	C	2	GAL	C1-C2-C3	-2.66	106.40	109.67
2	C	1	NAG	O1-C1-O5	-2.60	102.58	110.38
2	D	2	GAL	O5-C1-C2	-2.57	106.81	110.77
2	D	2	GAL	C3-C4-C5	-2.35	106.05	110.24
2	C	1	NAG	C6-C5-C4	2.31	118.42	113.00

There are no chirality outliers.

All (4) torsion outliers are listed below:

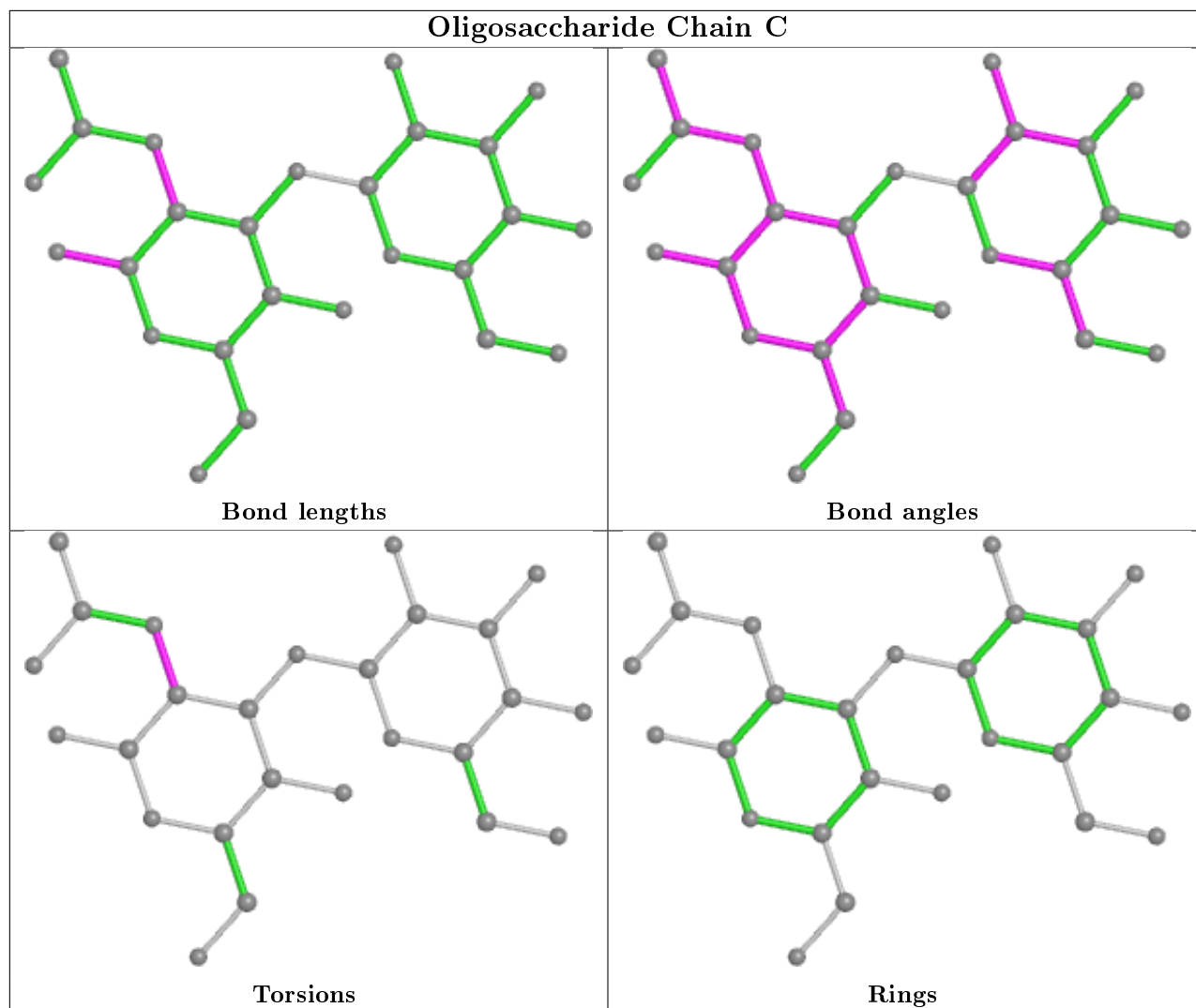
Mol	Chain	Res	Type	Atoms
2	C	1	NAG	C1-C2-N2-C7
2	D	1	NAG	C1-C2-N2-C7
2	D	1	NAG	C3-C2-N2-C7
2	C	1	NAG	C3-C2-N2-C7

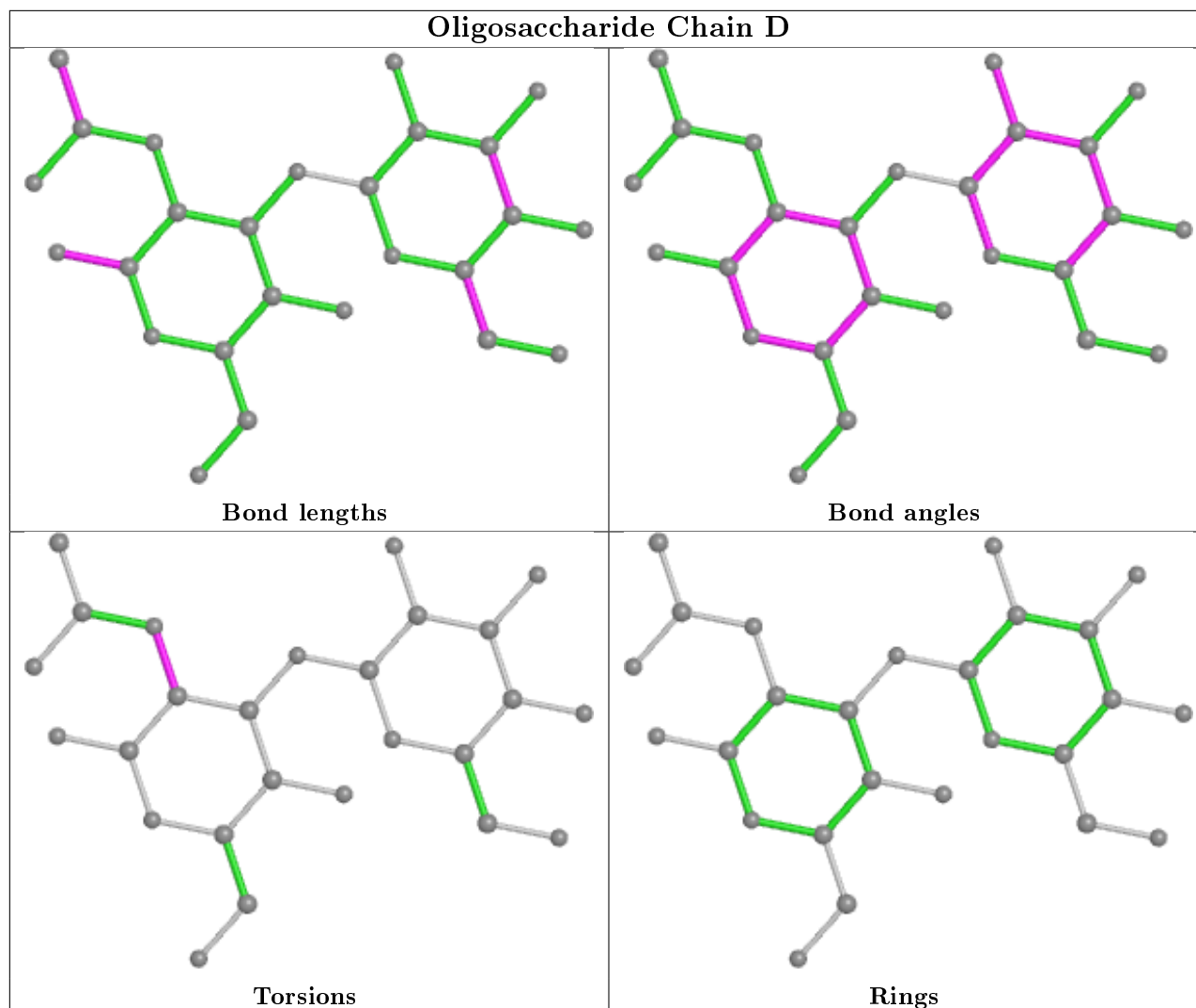
There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	NAG	2	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](#)

8 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	A	706	-	4,4,4	1.24	1 (25%)	6,6,6	0.69	0
3	SO4	B	706	-	4,4,4	0.64	0	6,6,6	0.94	0
3	SO4	A	704	-	4,4,4	0.59	0	6,6,6	2.00	2 (33%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	SO4	B	705	-	4,4,4	1.10	0	6,6,6	1.05	0
3	SO4	A	703	-	4,4,4	0.50	0	6,6,6	0.77	0
3	SO4	B	703	-	4,4,4	0.42	0	6,6,6	0.59	0
3	SO4	B	704	-	4,4,4	0.79	0	6,6,6	1.33	0
3	SO4	A	705	-	4,4,4	0.32	0	6,6,6	0.98	0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	706	SO4	O1-S	2.15	1.57	1.46

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	704	SO4	O4-S-O1	-3.46	91.23	109.31
3	A	704	SO4	O3-S-O2	2.11	120.33	109.31

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 [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	633/644 (98%)	-0.49	13 (2%) 63 59	7, 14, 30, 76	0
1	B	632/644 (98%)	-0.46	11 (1%) 70 66	6, 13, 33, 84	0
All	All	1265/1288 (98%)	-0.47	24 (1%) 66 63	6, 14, 32, 84	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	616	LYS	9.4
1	A	616	LYS	5.6
1	B	108	SER	5.4
1	B	617	ASP	5.0
1	A	617	ASP	4.4
1	B	107	GLY	4.2
1	A	662	SER	4.2
1	A	397	ALA	4.1
1	A	398	GLY	4.0
1	B	379	ASN	3.9
1	A	30	SER	3.6
1	B	618	GLY	3.3
1	A	379	ASN	3.2
1	A	618	GLY	3.0
1	B	30	SER	2.9
1	B	109	ALA	2.7
1	A	615	ASP	2.5
1	B	224	ASP	2.5
1	A	399	ARG	2.3
1	B	615	ASP	2.3
1	B	406	GLN	2.2
1	A	381	LYS	2.2
1	A	378	VAL	2.1
1	A	224	ASP	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.

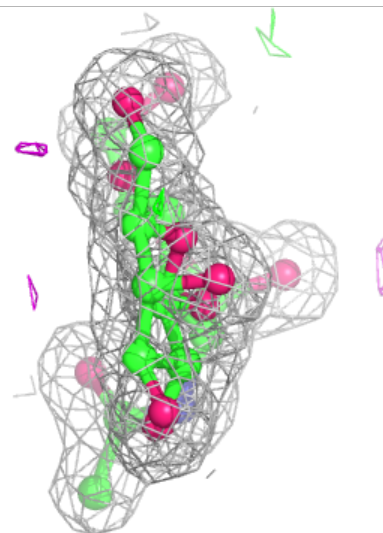
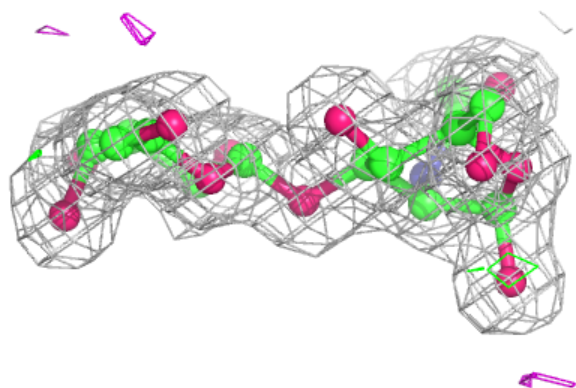
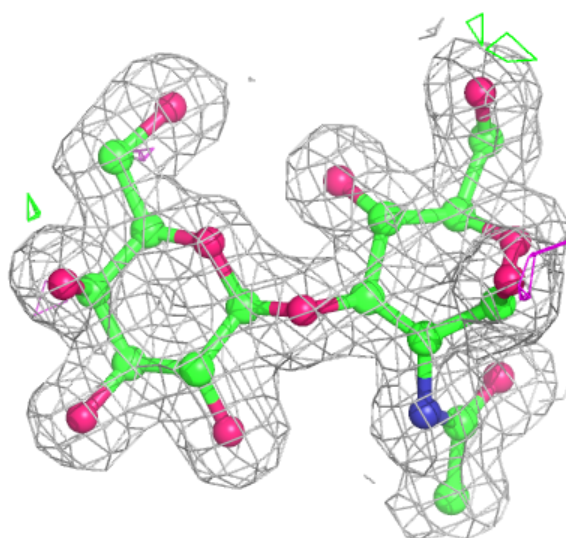
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	GAL	C	2	11/12	0.97	0.07	10,11,12,12	0
2	GAL	D	2	11/12	0.98	0.08	9,9,10,11	0
2	NAG	D	1	15/15	0.99	0.07	8,9,11,12	0
2	NAG	C	1	15/15	0.99	0.06	8,11,12,14	0

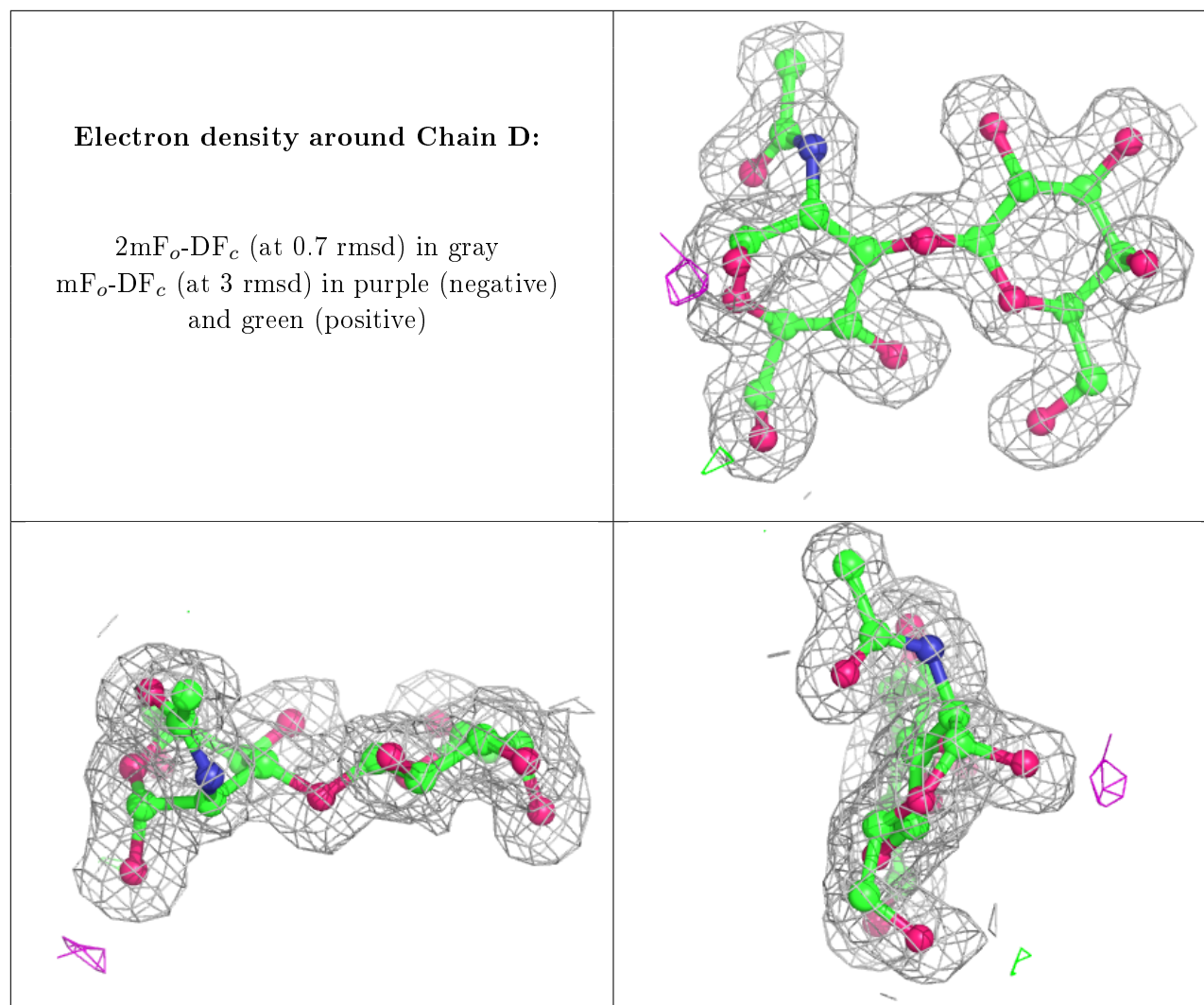
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**Electron density around Chain C:**

$2mF_o-DF_c$  (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	SO4	B	705	5/5	0.90	0.15	24,30,38,42	0
3	SO4	A	706	5/5	0.93	0.20	27,34,40,48	0
3	SO4	B	706	5/5	0.94	0.14	41,54,55,60	0
3	SO4	B	704	5/5	0.95	0.16	31,36,44,45	0
3	SO4	A	705	5/5	0.95	0.13	42,44,49,54	0
3	SO4	A	704	5/5	0.97	0.14	21,25,27,28	0
3	SO4	A	703	5/5	0.99	0.06	11,11,12,12	0
3	SO4	B	703	5/5	1.00	0.05	9,9,10,10	0

## 6.5 Other polymers

There are no such residues in this entry.