



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 6, 2025 – 04:26 PM JST

PDB ID : 8WNA
Title : Crystal structure of GH1 beta-glucosidase Td2F2 E352Q laminaribiose complex
Authors : Lin, C.C.; Lu, T.J.; Fushinobu, S.
Deposited on : 2023-10-05
Resolution : 1.16 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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

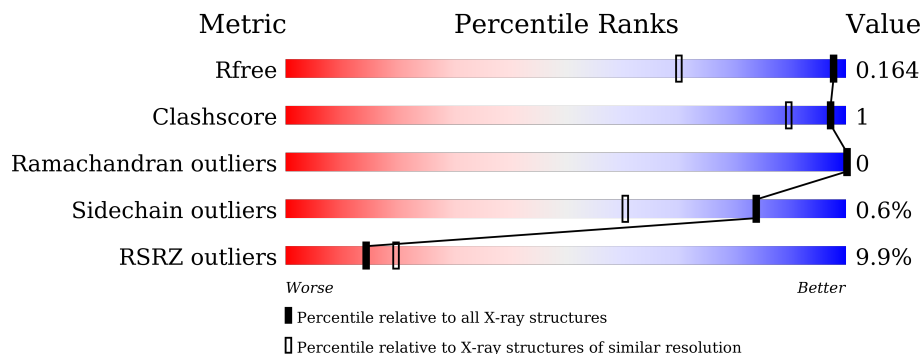
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.16 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}	164625	1049 (1.18-1.14)
Clashscore	180529	1146 (1.18-1.14)
Ramachandran outliers	177936	1120 (1.18-1.14)
Sidechain outliers	177891	1120 (1.18-1.14)
RSRZ outliers	164620	1049 (1.18-1.14)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 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	465	 9% 90% 5% 5%
2	C	2	 100%
2	D	2	 100%
3	B	2	 50% 50%

2 Entry composition [i](#)

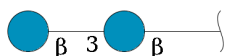
There are 7 unique types of molecules in this entry. The entry contains 3991 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 beta-glucosidase Td2F2.

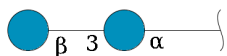
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	443	3553	2263	621	655	14	0	10	0

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



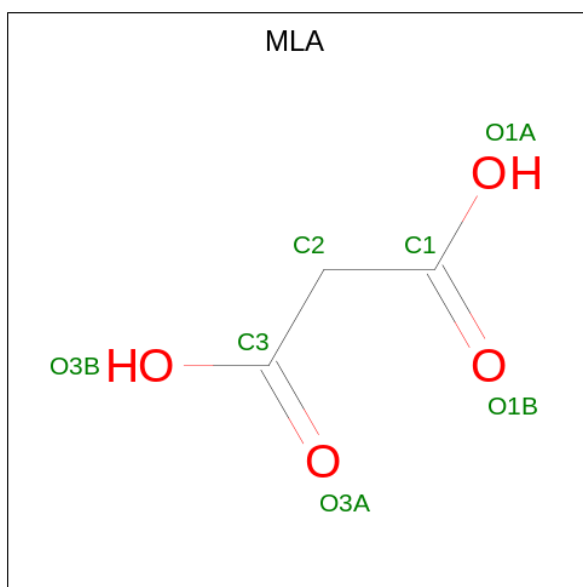
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
2	C	2	23	12	11	0	0	0
2	D	2	23	12	11	0	0	0

- Molecule 3 is an oligosaccharide called beta-D-glucopyranose-(1-3)-alpha-D-glucopyranose.



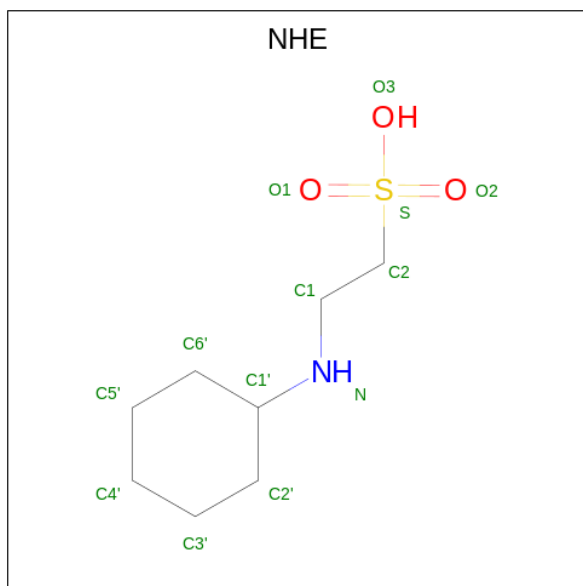
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
3	B	2	23	12	11	0	0	0

- Molecule 4 is MALONIC ACID (three-letter code: MLA) (formula: C₃H₄O₄).



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

- Molecule 5 is 2-[N-CYCLOHEXYLAMINO]ETHANE SULFONIC ACID (three-letter code: NHE) (formula: $C_8H_{17}NO_3S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	S	0	0
			13	8	1	3	1		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	O	S	0	0
			5	4	1		

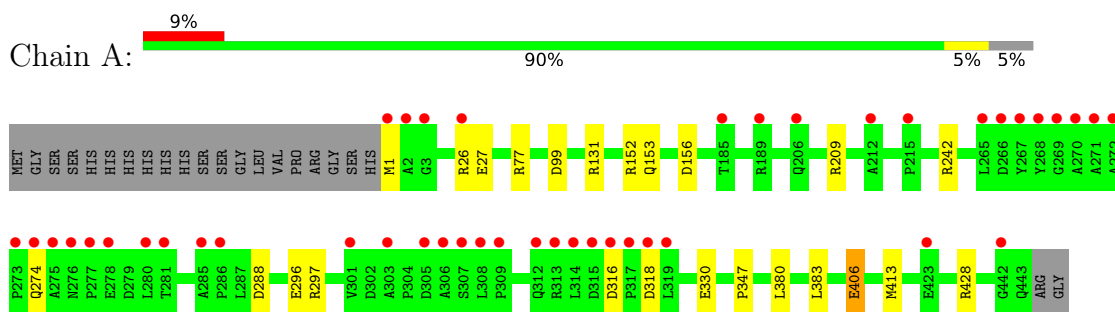
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	344	Total	O	0	0
			344	344		

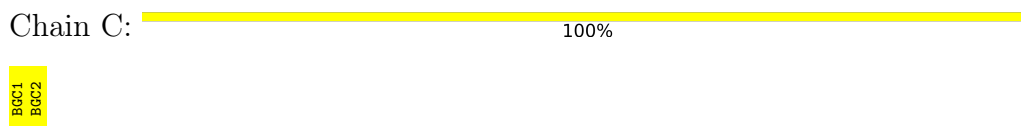
3 Residue-property plots [i](#)

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: beta-glucosidase Td2F2



- Molecule 2: beta-D-glucopyranose-(1-3)-beta-D-glucopyranose



- Molecule 2: beta-D-glucopyranose-(1-3)-beta-D-glucopyranose



- Molecule 3: beta-D-glucopyranose-(1-3)-alpha-D-glucopyranose



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	68.41Å 68.42Å 95.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.42 – 1.16 48.42 – 1.16	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.42-1.16) 99.9 (48.42-1.16)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.26 (at 1.16Å)	Xtrriage
Refinement program	REFMAC 5.8.0405	Depositor
R, R_{free}	0.151 , 0.164 0.150 , 0.164	Depositor DCC
R_{free} test set	7437 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å ²)	9.5	Xtrriage
Anisotropy	0.008	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 30.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.013 for k,h,-l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3991	wwPDB-VP
Average B, all atoms (Å ²)	12.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GLC, NHE, BGC, SO4, MLA

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	0.81	5/3660 (0.1%)	1.01	12/4994 (0.2%)

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	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	27	GLU	CD-OE1	-7.54	1.17	1.25
1	A	406[A]	GLU	CD-OE1	-6.15	1.18	1.25
1	A	406[B]	GLU	CD-OE1	-6.15	1.18	1.25
1	A	330	GLU	CD-OE1	-5.61	1.19	1.25
1	A	296	GLU	CD-OE1	5.19	1.31	1.25

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	209[A]	ARG	NE-CZ-NH2	-12.41	114.09	120.30
1	A	209[B]	ARG	NE-CZ-NH2	-12.41	114.09	120.30
1	A	131	ARG	NE-CZ-NH1	9.04	124.82	120.30
1	A	242	ARG	NE-CZ-NH1	-7.24	116.68	120.30
1	A	209[A]	ARG	NE-CZ-NH1	6.60	123.60	120.30
1	A	209[B]	ARG	NE-CZ-NH1	6.60	123.60	120.30
1	A	152	ARG	NE-CZ-NH2	-6.52	117.04	120.30
1	A	131	ARG	NE-CZ-NH2	-6.05	117.28	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	297	ARG	NE-CZ-NH1	-5.91	117.35	120.30
1	A	99	ASP	CB-CG-OD2	-5.59	113.27	118.30
1	A	428	ARG	NE-CZ-NH2	5.46	123.03	120.30
1	A	156	ASP	CB-CG-OD1	5.28	123.05	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	77	ARG	Sidechain

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	3553	0	3377	6	0
2	C	23	0	21	0	0
2	D	23	0	21	0	0
3	B	23	0	20	0	0
4	A	7	0	2	0	0
5	A	13	0	16	0	0
6	A	5	0	0	0	0
7	A	344	0	0	1	0
All	All	3991	0	3457	6	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 (6) 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:406[B]:GLU:HG3	1:A:413:MET:O	2.01	0.61
1:A:316:ASP:OD1	1:A:318:ASP:HB2	2.13	0.48
1:A:26:ARG:HA	1:A:26:ARG:HD2	1.79	0.46
1:A:380:LEU:HD23	1:A:383:LEU:HD12	1.99	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:153[A]:GLN:HG3	7:A:777:HOH:O	2.18	0.44
1:A:288:ASP:O	1:A:347:PRO:HD2	2.20	0.41

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	451/465 (97%)	437 (97%)	14 (3%)	0	100 100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	354/363 (98%)	352 (99%)	2 (1%)	84 60

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	274	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	262	GLN

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](#)

6 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$
3	GLC	B	1	3	12,12,12	0.89	0	17,17,17	0.91	0
3	BGC	B	2	3	11,11,12	0.87	1 (9%)	15,15,17	0.68	0
2	BGC	C	1	2	12,12,12	0.39	0	17,17,17	2.59	7 (41%)
2	BGC	C	2	2	11,11,12	0.99	0	15,15,17	0.86	1 (6%)
2	BGC	D	1	2	12,12,12	0.79	0	17,17,17	1.80	4 (23%)
2	BGC	D	2	2	11,11,12	1.54	2 (18%)	15,15,17	1.61	4 (26%)

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
3	GLC	B	1	3	-	0/2/22/22	0/1/1/1
3	BGC	B	2	3	-	0/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	BGC	C	1	2	-	0/2/22/22	0/1/1/1
2	BGC	C	2	2	-	0/2/19/22	0/1/1/1
2	BGC	D	1	2	-	0/2/22/22	0/1/1/1
2	BGC	D	2	2	-	2/2/19/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	2	BGC	O5-C1	-3.20	1.38	1.43
2	D	2	BGC	O4-C4	-2.21	1.37	1.43
3	B	2	BGC	C2-C3	-2.03	1.49	1.52

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1	BGC	C4-C3-C2	-5.16	101.82	110.82
2	C	1	BGC	O2-C2-C3	4.61	121.02	110.35
2	D	1	BGC	O5-C5-C4	-4.10	102.25	109.69
2	C	1	BGC	C3-C4-C5	-3.65	103.73	110.24
2	C	1	BGC	O4-C4-C3	3.55	118.56	110.35
2	C	1	BGC	O3-C3-C4	3.53	118.50	110.35
2	D	1	BGC	O3-C3-C2	3.45	118.32	110.35
2	D	2	BGC	O2-C2-C1	3.40	116.10	109.15
2	C	1	BGC	C1-C2-C3	-2.96	104.16	110.31
2	C	1	BGC	O3-C3-C2	2.87	116.97	110.35
2	D	1	BGC	O1-C1-C2	2.82	116.98	109.03
2	D	1	BGC	O3-C3-C4	-2.57	104.40	110.35
2	D	2	BGC	C6-C5-C4	2.50	118.86	113.00
2	D	2	BGC	C1-C2-C3	2.25	112.43	109.67
2	D	2	BGC	O3-C3-C4	2.12	115.25	110.35
2	C	2	BGC	O5-C1-C2	-2.08	107.56	110.77

There are no chirality outliers.

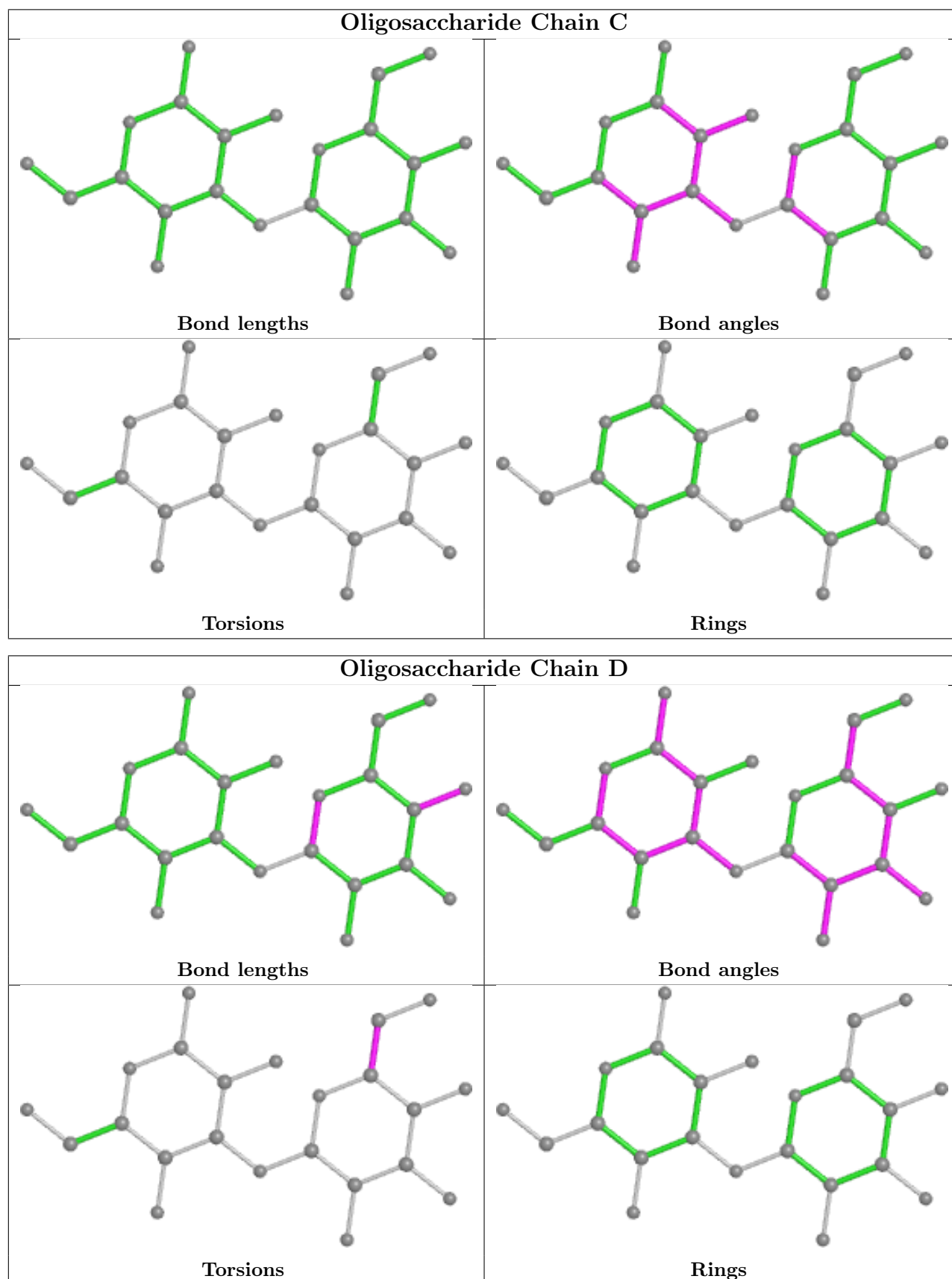
All (2) torsion outliers are listed below:

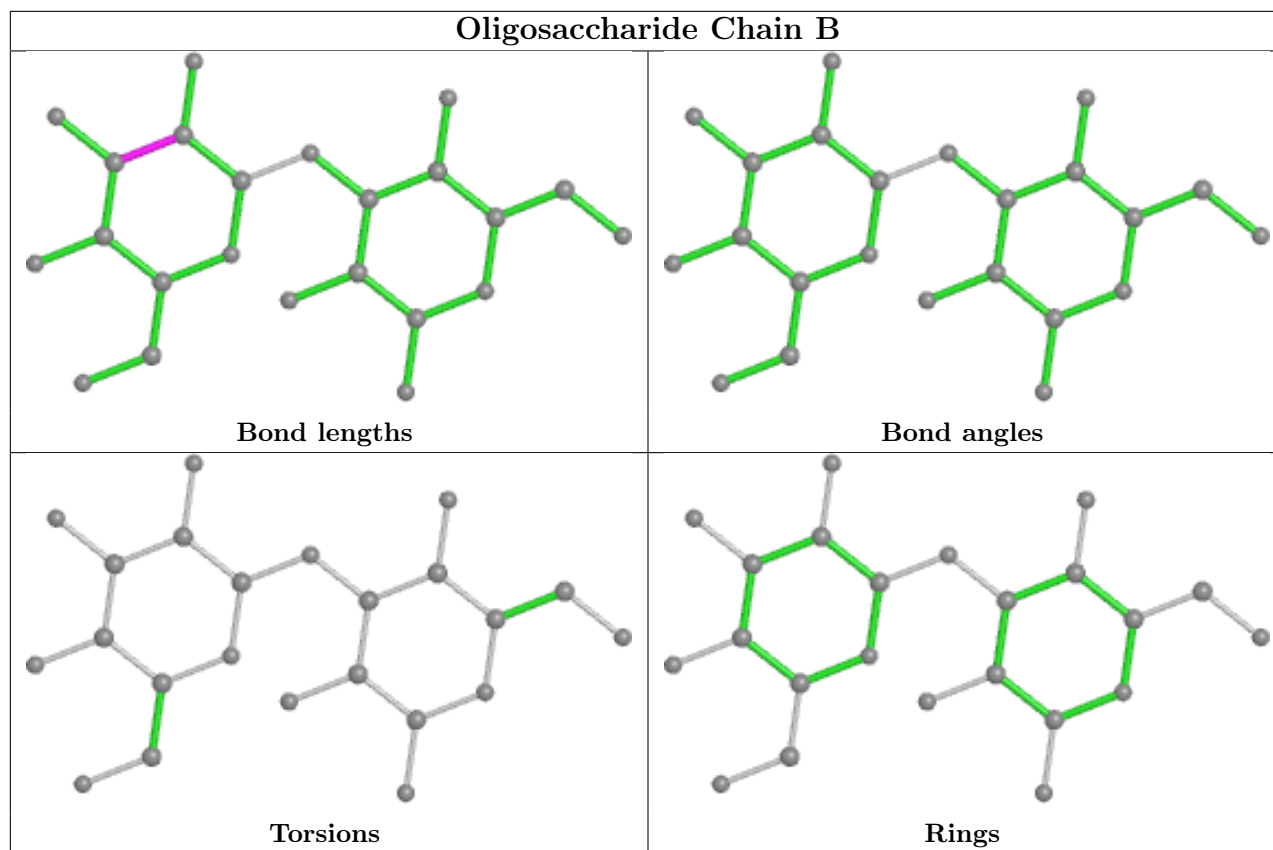
Mol	Chain	Res	Type	Atoms
2	D	2	BGC	O5-C5-C6-O6
2	D	2	BGC	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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





5.6 Ligand geometry [i](#)

3 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	$\# Z > 2$
6	SO4	A	503	-	4,4,4	0.29	0	6,6,6	0.19	0
4	MLA	A	501	-	6,6,6	1.13	0	7,7,7	0.67	0
5	NHE	A	502	-	13,13,13	0.38	0	16,17,17	0.78	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	MLA	A	501	-	-	3/4/4/4	-
5	NHE	A	502	-	-	0/7/15/15	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	501	MLA	C1-C2-C3-O3B
4	A	501	MLA	C1-C2-C3-O3A
4	A	501	MLA	O1A-C1-C2-C3

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, 95th 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 ²)	Q<0.9
1	A	443/465 (95%)	0.40	44 (9%) 14 20	4, 9, 24, 61	10 (2%)

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	317	PRO	7.1
1	A	270	ALA	6.2
1	A	275	ALA	5.2
1	A	318	ASP	4.9
1	A	308	LEU	4.6
1	A	315	ASP	4.5
1	A	274	GLN	4.1
1	A	269	GLY	4.0
1	A	272	ALA	3.8
1	A	277	PRO	3.8
1	A	1	MET	3.7
1	A	306	ALA	3.6
1	A	276	ASN	3.6
1	A	271	ALA	3.5
1	A	266	ASP	3.2
1	A	280	LEU	3.1
1	A	319	LEU	3.1
1	A	278	GLU	3.1
1	A	303	ALA	3.1
1	A	26	ARG	3.1
1	A	314	LEU	3.0
1	A	316	ASP	2.9
1	A	285	ALA	2.9
1	A	281	THR	2.8
1	A	309	PRO	2.8
1	A	312	GLN	2.8
1	A	307	SER	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	313	ARG	2.7
1	A	206	GLN	2.7
1	A	2	ALA	2.7
1	A	268	TYR	2.6
1	A	286	PRO	2.6
1	A	442	GLY	2.6
1	A	305	ASP	2.5
1	A	423	GLU	2.4
1	A	273	PRO	2.4
1	A	189	ARG	2.4
1	A	265	LEU	2.3
1	A	267	TYR	2.3
1	A	301	VAL	2.2
1	A	3	GLY	2.1
1	A	212	ALA	2.1
1	A	185	THR	2.1
1	A	215	PRO	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](#)

SUGAR-RSR INFOmissingINFO

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, 95th 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 ²)	Q<0.9
6	SO4	A	503	5/5	0.75	0.18	45,45,67,72	0
4	MLA	A	501	7/7	0.93	0.09	13,18,21,22	0
5	NHE	A	502	13/13	0.98	0.06	8,9,12,12	0

6.5 Other polymers [i](#)

There are no such residues in this entry.