



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 06:16 PM UTC

PDB ID : 4U5K / pdb_00004u5k
Title : Complex structure of mutant CtCel5E (E314A) with cellobiose
Authors : Guo, R.T.; Huang, C.H.; Wu, T.H.
Deposited on : 2014-07-25
Resolution : 2.65 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

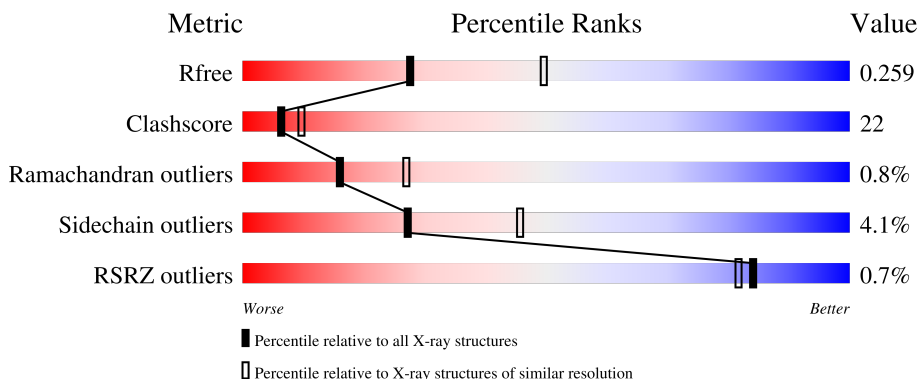
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.65 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}	180053	1110 (2.66-2.66)
Clashscore	190562	1141 (2.66-2.66)
Ramachandran outliers	187476	1126 (2.66-2.66)
Sidechain outliers	187428	1126 (2.66-2.66)
RSRZ outliers	180081	1110 (2.66-2.66)

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	403	
1	B	403	
2	C	2	
2	D	2	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5575 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 Endoglucanase H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	304	Total	C	N	O	S	0	0	0
			2532	1625	424	472	11			
1	B	298	Total	C	N	O	S	0	0	0
			2473	1588	411	463	11			

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP P16218
A	2	GLY	-	expression tag	UNP P16218
A	3	SER	-	expression tag	UNP P16218
A	4	SER	-	expression tag	UNP P16218
A	5	HIS	-	expression tag	UNP P16218
A	6	HIS	-	expression tag	UNP P16218
A	7	HIS	-	expression tag	UNP P16218
A	8	HIS	-	expression tag	UNP P16218
A	9	HIS	-	expression tag	UNP P16218
A	10	HIS	-	expression tag	UNP P16218
A	11	SER	-	expression tag	UNP P16218
A	12	SER	-	expression tag	UNP P16218
A	13	GLY	-	expression tag	UNP P16218
A	14	LEU	-	expression tag	UNP P16218
A	15	VAL	-	expression tag	UNP P16218
A	16	PRO	-	expression tag	UNP P16218
A	17	ARG	-	expression tag	UNP P16218
A	18	GLY	-	expression tag	UNP P16218
A	19	SER	-	expression tag	UNP P16218
A	20	HIS	-	expression tag	UNP P16218
A	21	MET	-	expression tag	UNP P16218
A	22	ALA	-	expression tag	UNP P16218
A	23	SER	-	expression tag	UNP P16218
A	24	MET	-	expression tag	UNP P16218
A	25	THR	-	expression tag	UNP P16218

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	GLY	-	expression tag	UNP P16218
A	27	GLY	-	expression tag	UNP P16218
A	28	GLN	-	expression tag	UNP P16218
A	29	GLN	-	expression tag	UNP P16218
A	30	MET	-	expression tag	UNP P16218
A	31	GLY	-	expression tag	UNP P16218
A	32	ARG	-	expression tag	UNP P16218
A	33	ILE	-	expression tag	UNP P16218
A	34	GLU	-	expression tag	UNP P16218
A	35	GLY	-	expression tag	UNP P16218
A	36	ARG	-	expression tag	UNP P16218
A	37	GLU	-	expression tag	UNP P16218
A	38	PHE	-	expression tag	UNP P16218
A	314	ALA	GLU	engineered mutation	UNP P16218
B	1	MET	-	expression tag	UNP P16218
B	2	GLY	-	expression tag	UNP P16218
B	3	SER	-	expression tag	UNP P16218
B	4	SER	-	expression tag	UNP P16218
B	5	HIS	-	expression tag	UNP P16218
B	6	HIS	-	expression tag	UNP P16218
B	7	HIS	-	expression tag	UNP P16218
B	8	HIS	-	expression tag	UNP P16218
B	9	HIS	-	expression tag	UNP P16218
B	10	HIS	-	expression tag	UNP P16218
B	11	SER	-	expression tag	UNP P16218
B	12	SER	-	expression tag	UNP P16218
B	13	GLY	-	expression tag	UNP P16218
B	14	LEU	-	expression tag	UNP P16218
B	15	VAL	-	expression tag	UNP P16218
B	16	PRO	-	expression tag	UNP P16218
B	17	ARG	-	expression tag	UNP P16218
B	18	GLY	-	expression tag	UNP P16218
B	19	SER	-	expression tag	UNP P16218
B	20	HIS	-	expression tag	UNP P16218
B	21	MET	-	expression tag	UNP P16218
B	22	ALA	-	expression tag	UNP P16218
B	23	SER	-	expression tag	UNP P16218
B	24	MET	-	expression tag	UNP P16218
B	25	THR	-	expression tag	UNP P16218
B	26	GLY	-	expression tag	UNP P16218
B	27	GLY	-	expression tag	UNP P16218
B	28	GLN	-	expression tag	UNP P16218

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	29	GLN	-	expression tag	UNP P16218
B	30	MET	-	expression tag	UNP P16218
B	31	GLY	-	expression tag	UNP P16218
B	32	ARG	-	expression tag	UNP P16218
B	33	ILE	-	expression tag	UNP P16218
B	34	GLU	-	expression tag	UNP P16218
B	35	GLY	-	expression tag	UNP P16218
B	36	ARG	-	expression tag	UNP P16218
B	37	GLU	-	expression tag	UNP P16218
B	38	PHE	-	expression tag	UNP P16218
B	314	ALA	GLU	engineered mutation	UNP P16218

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



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	284	Total	O	0	0
			284	284		
3	B	240	Total	O	0	0
			240	240		

SER
TYR
SER
PRO
LYS
PRO
SER
PRO
THR
PRO
ARG
PRO
THR
LYS
PRO
PRO
VAL
THR
PRO

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



BGC1
BGC2

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



BGC1
BGC2

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	74.94Å 74.94Å 254.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.65 25.00 – 2.65	Depositor EDS
% Data completeness (in resolution range)	94.0 (25.00-2.65) 94.0 (25.00-2.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.24 (at 2.64Å)	Xtriage
Refinement program	CNS 1.21	Depositor
R, R_{free}	0.183 , 0.255 0.185 , 0.259	Depositor DCC
R_{free} test set	1000 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å ²)	39.1	Xtriage
Anisotropy	0.080	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 55.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	5575	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BGC

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	0.43	0/2605	1.00	10/3533 (0.3%)
1	B	0.39	0/2542	0.91	9/3446 (0.3%)
All	All	0.41	0/5147	0.96	19/6979 (0.3%)

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	162	VAL	N-CA-C	-10.95	95.61	109.30
1	A	283	TRP	N-CA-C	9.32	122.55	110.53
1	A	162	VAL	N-CA-C	-9.00	96.96	108.84
1	A	209	GLU	CA-C-N	8.58	128.57	119.05
1	A	209	GLU	C-N-CA	8.58	128.57	119.05

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	2532	0	2392	104	0
1	B	2473	0	2337	119	0
2	C	23	0	20	2	0
2	D	23	0	20	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	284	0	0	7	0
3	B	240	0	0	2	0
All	All	5575	0	4769	217	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 22.

The worst 5 of 217 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:82:VAL:HG21	1:A:309:PRO:HB2	1.26	1.17
1:B:82:VAL:HG21	1:B:309:PRO:HB2	1.41	1.01
1:B:134:MET:HE1	1:B:138:PRO:HD2	1.54	0.90
1:B:352:PHE:HA	1:B:360:ALA:HB3	1.54	0.88
1:B:315:PHE:CZ	1:B:333:ILE:HG22	2.10	0.86

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	300/403 (74%)	275 (92%)	22 (7%)	3 (1%)	12	20
1	B	292/403 (72%)	266 (91%)	24 (8%)	2 (1%)	18	30
All	All	592/806 (73%)	541 (91%)	46 (8%)	5 (1%)	16	27

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	357	ASN
1	A	358	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	177	TYR
1	B	196	LYS
1	B	316	ALA

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	270/353 (76%)	260 (96%)	10 (4%)	30	50
1	B	264/353 (75%)	252 (96%)	12 (4%)	24	42
All	All	534/706 (76%)	512 (96%)	22 (4%)	27	46

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	203	LEU
1	B	286	GLN
1	B	254	ASN
1	B	293	VAL
1	A	339	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	180	ASN
1	B	197	ASN
1	B	357	ASN
1	B	268	HIS
1	B	286	GLN

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 ⓘ

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	BGC	C	1	2	12,12,12	3.08	3 (25%)	17,17,17	1.16	1 (5%)
2	BGC	C	2	2	11,11,12	1.69	4 (36%)	15,15,17	0.78	0
2	BGC	D	1	2	12,12,12	3.15	3 (25%)	17,17,17	1.15	1 (5%)
2	BGC	D	2	2	11,11,12	1.58	2 (18%)	15,15,17	0.85	0

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

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

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1	BGC	O6-C6	-9.48	1.02	1.42
2	C	1	BGC	O6-C6	-9.47	1.02	1.42
2	D	2	BGC	O5-C1	3.39	1.49	1.43
2	D	1	BGC	C4-C5	3.36	1.60	1.53
2	C	2	BGC	O5-C1	2.95	1.48	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1	BGC	O6-C6-C5	3.21	122.28	111.33
2	D	1	BGC	O6-C6-C5	3.13	122.00	111.33

There are no chirality outliers.

All (2) torsion outliers are listed below:

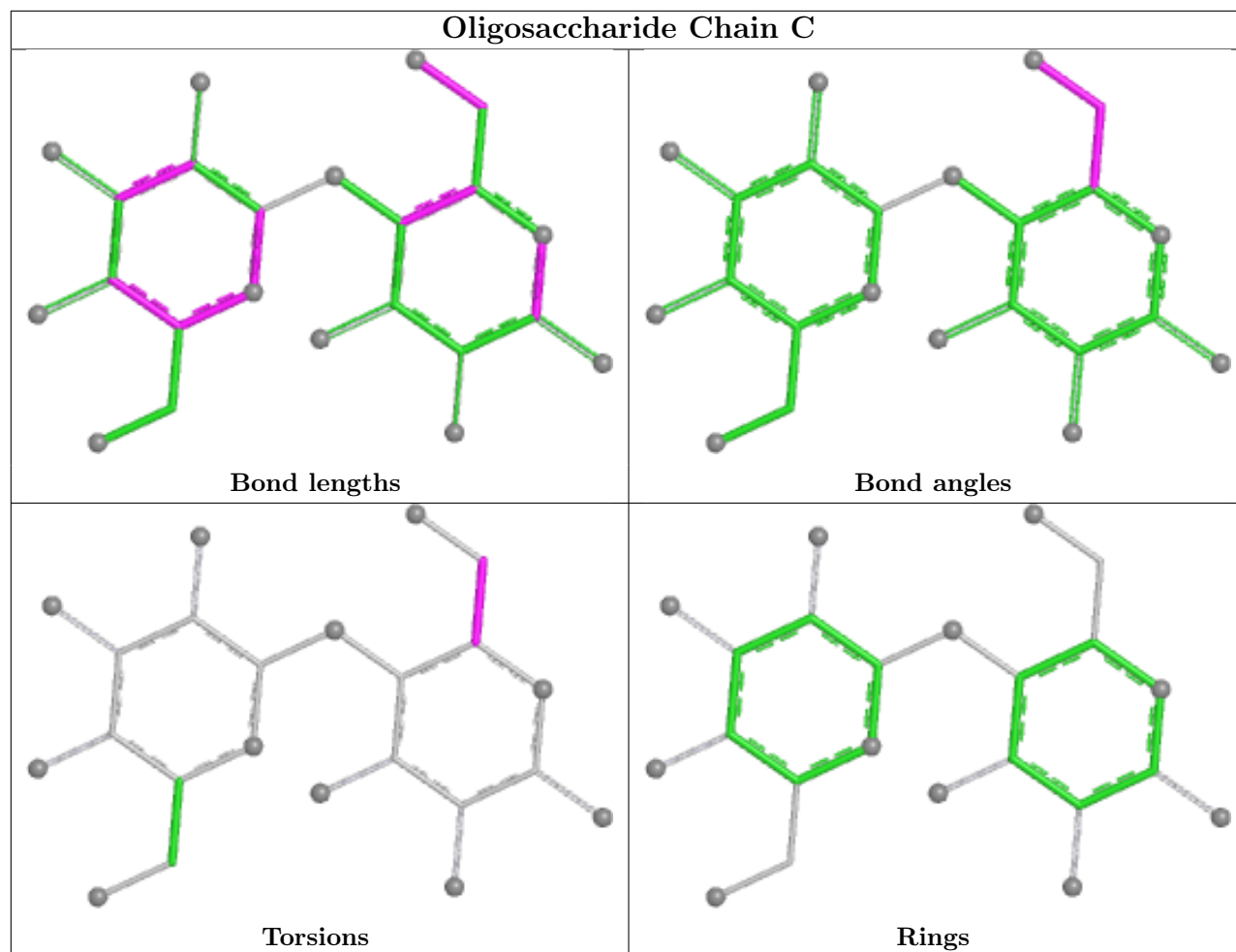
Mol	Chain	Res	Type	Atoms
2	C	1	BGC	C4-C5-C6-O6
2	C	1	BGC	O5-C5-C6-O6

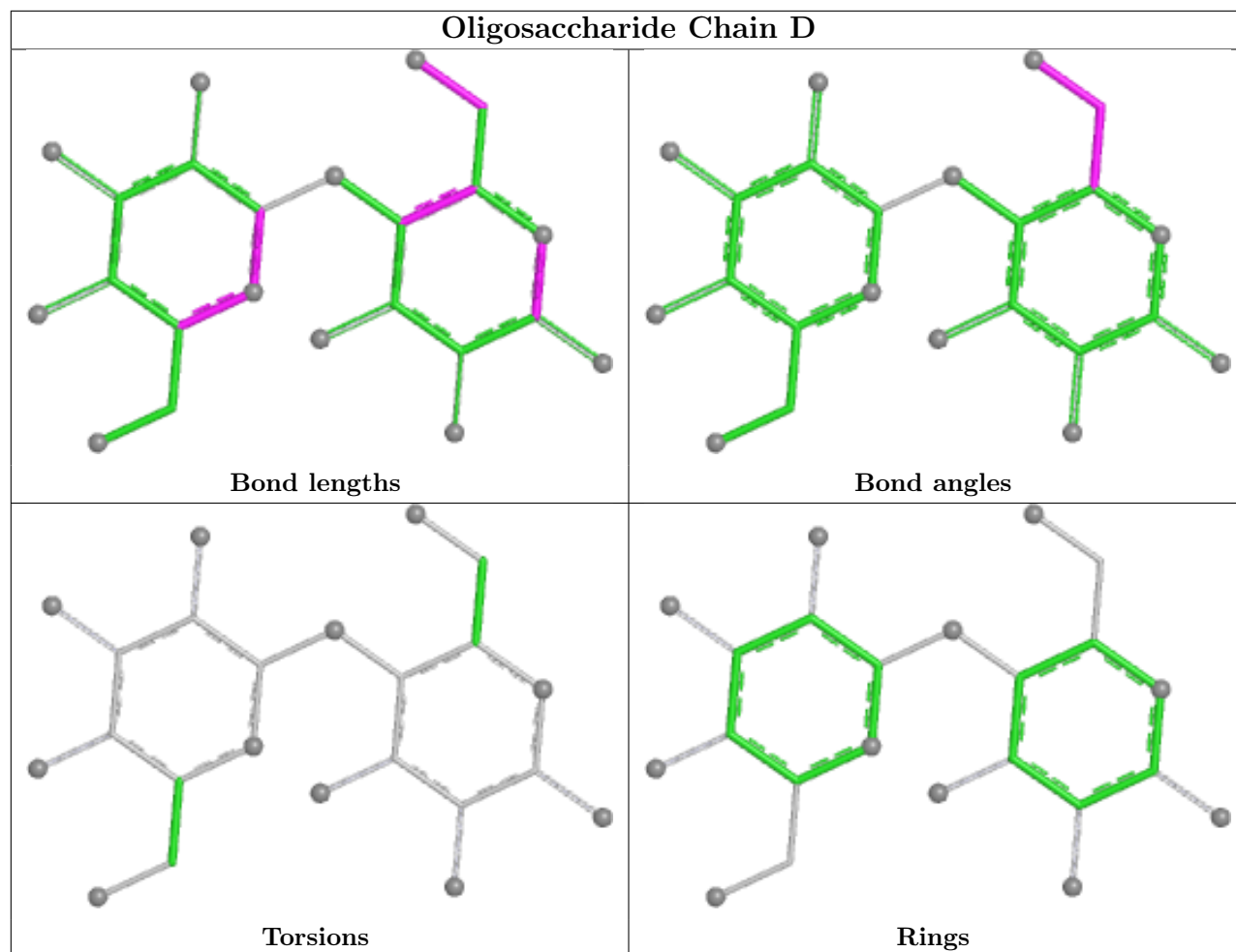
There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1	BGC	2	0
2	C	1	BGC	1	0
2	C	2	BGC	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](#)

There are no ligands in this entry.

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 [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, 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	304/403 (75%)	-0.50	0 100 100	16, 32, 64, 103	0
1	B	298/403 (73%)	-0.35	4 (1%) 75 71	22, 40, 78, 100	0
All	All	602/806 (74%)	-0.42	4 (0%) 84 82	16, 35, 74, 103	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	353	GLY	2.5
1	B	275	PHE	2.4
1	B	283	TRP	2.1
1	B	281	GLY	2.1

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, 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.

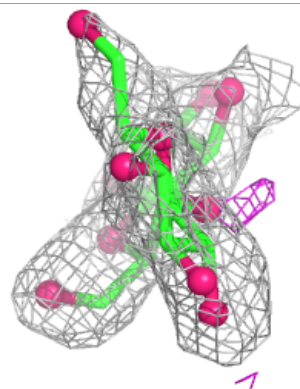
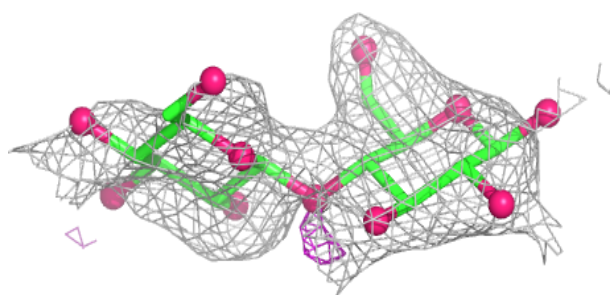
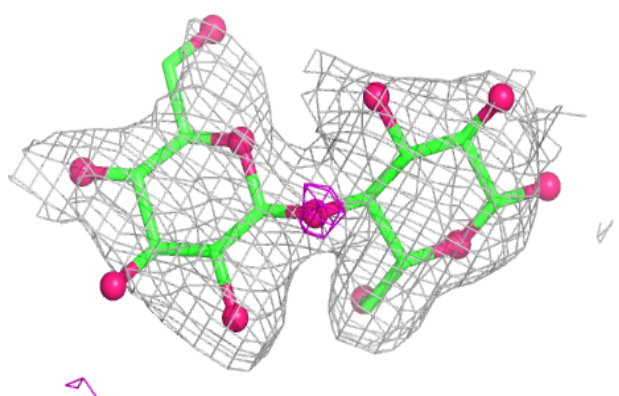
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	BGC	D	1	12/12	0.78	0.17	100,101,102,102	0
2	BGC	D	2	11/12	0.78	0.18	96,98,98,99	0
2	BGC	C	2	11/12	0.83	0.11	64,68,70,70	0
2	BGC	C	1	12/12	0.85	0.12	57,62,65,66	0

The following is a graphical depiction of the model fit to experimental electron density for oligosac-

charide. Each fit is shown from different orientation to approximate a three-dimensional view.

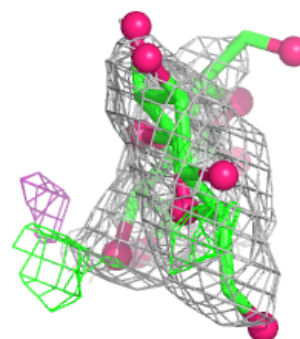
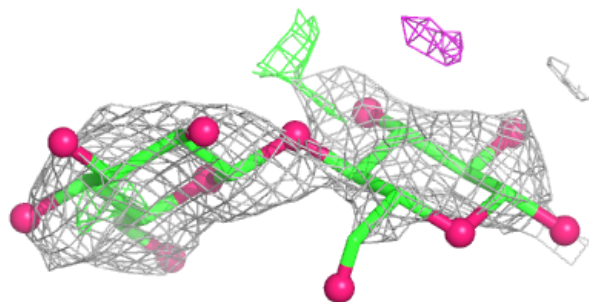
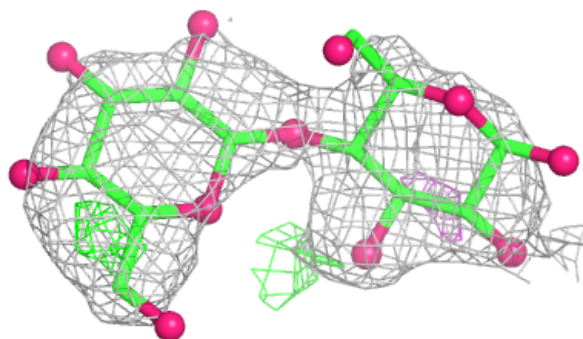
Electron density around Chain C:

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



Electron density around Chain D:

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



6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.