WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 4, 2024 – 02:07 pm GMT

PDB ID : 5CLT

Title : Crystal structure of human glycogen branching enzyme (GBE1) in complex with acarbose

Authors : Krojer, T.; Froese, D.S.; Goubin, S.; Strain-Damerell, C.; Mahajan, P.; Burgess-Brown, N.; von Delft, F.; Bountra, C.; Arrowsmith, C.H.; Edwards, A.; Yue, W.; Structural Genomics Consortium (SGC)

Deposited on : 2015-07-16

Resolution : 2.79 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 (i) 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 (i)) were used in the production of this report:
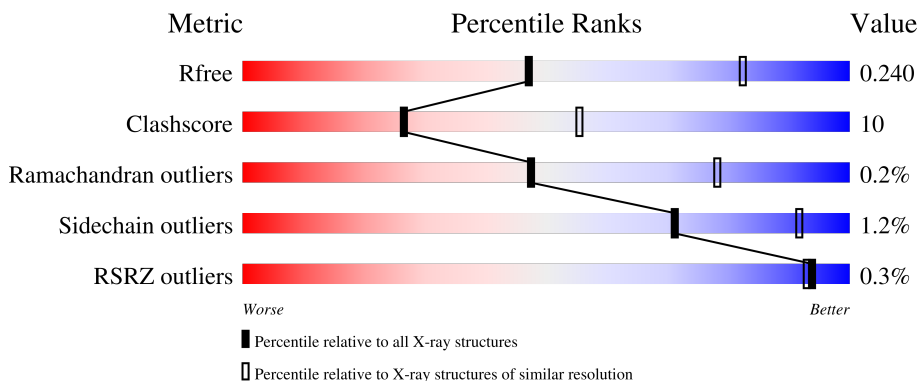
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.79 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 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 <=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	668	80% 16% ·
1	B	668	71% 24% 5%
1	C	668	% 76% 19% · 5%
2	D	3	33% 67%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	3	33% 67%
2	F	3	100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GLC	F	1	-	-	-	X

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 15585 atoms, of which 129 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 1,4-alpha-glucan-branching enzyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	646	Total	C	N	O	S	0	0	0
			5184	3335	879	947	23			
1	B	637	Total	C	N	O	S	0	0	0
			5064	3253	857	932	22			
1	C	637	Total	C	N	O	S	0	0	0
			5076	3263	860	930	23			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	33	TYR	-	expression tag	UNP Q04446
A	34	PHE	-	expression tag	UNP Q04446
A	35	GLN	-	expression tag	UNP Q04446
A	36	SER	-	expression tag	UNP Q04446
A	37	MET	-	expression tag	UNP Q04446
A	265	SER	THR	conflict	UNP Q04446
A	334	VAL	ILE	conflict	UNP Q04446
B	33	TYR	-	expression tag	UNP Q04446
B	34	PHE	-	expression tag	UNP Q04446
B	35	GLN	-	expression tag	UNP Q04446
B	36	SER	-	expression tag	UNP Q04446
B	37	MET	-	expression tag	UNP Q04446
B	265	SER	THR	conflict	UNP Q04446
B	334	VAL	ILE	conflict	UNP Q04446
C	33	TYR	-	expression tag	UNP Q04446
C	34	PHE	-	expression tag	UNP Q04446
C	35	GLN	-	expression tag	UNP Q04446
C	36	SER	-	expression tag	UNP Q04446
C	37	MET	-	expression tag	UNP Q04446
C	265	SER	THR	conflict	UNP Q04446
C	334	VAL	ILE	conflict	UNP Q04446

- Molecule 2 is an oligosaccharide called 4,6-dideoxy-4-{[(1S,4R,5S,6S)-4,5,6-trihydroxy-3-(hy
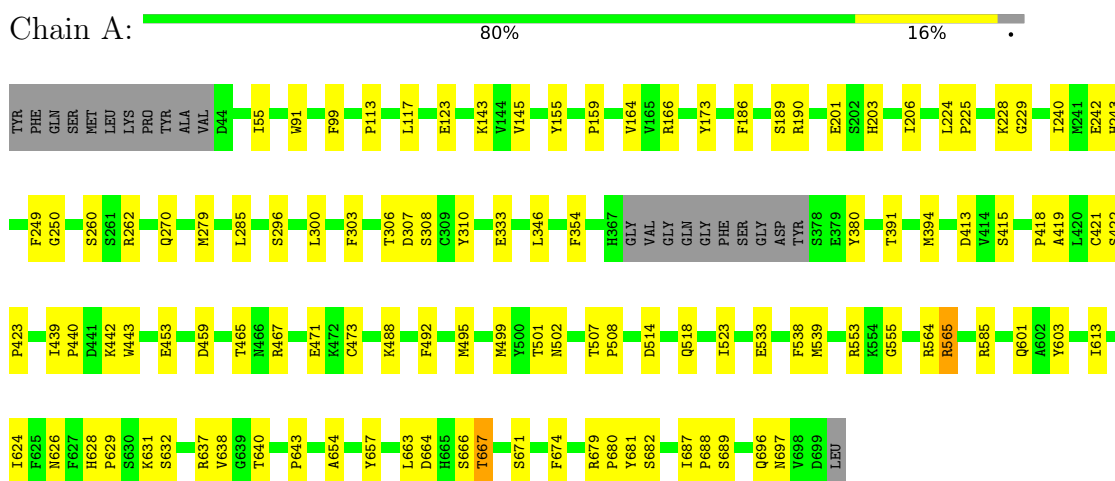
droxymethyl)cyclohex-2-en-1-yl]amino}-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.

| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	3	Total 87	C 25	H 43	N 1	O 18	0	0	0
2	E	3	Total 87	C 25	H 43	N 1	O 18	0	0	0
2	F	3	Total 87	C 25	H 43	N 1	O 18	0	0	0

3 Residue-property plots ⓘ

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.
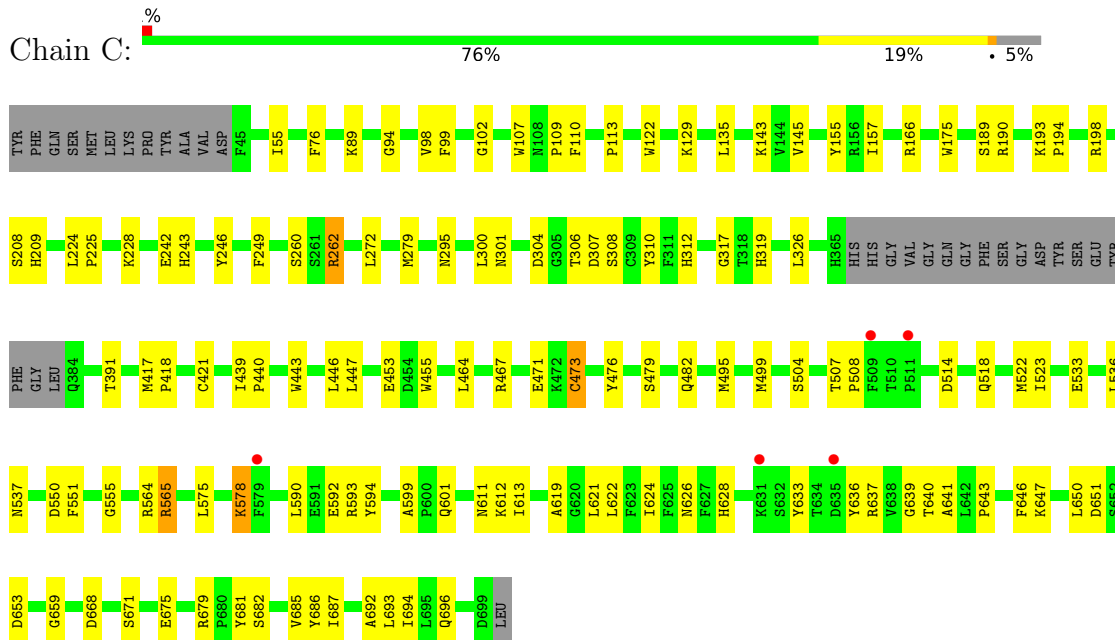
- Molecule 1: 1,4-alpha-glucan-branching enzyme
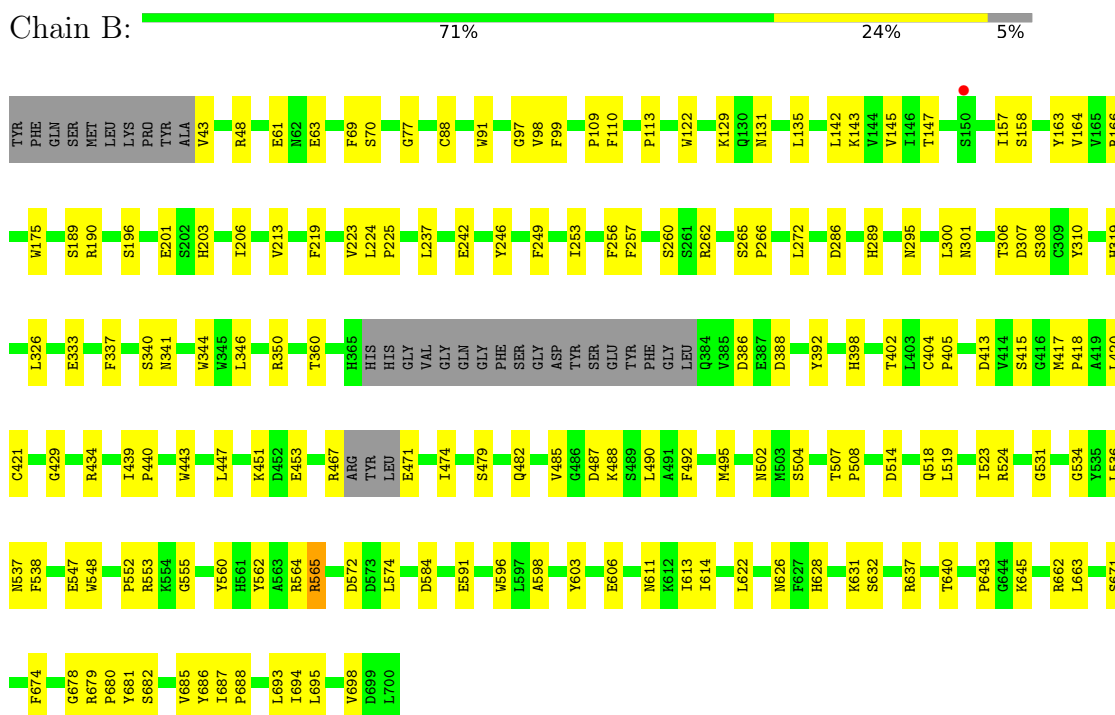
- Molecule 1: 1,4-alpha-glucan-branching enzyme

● Molecule 1: 1,4-alpha-glucan-branching enzyme

Chain C:

● Molecule 2: 4,6-dideoxy-4-{[(1S,4R,5S,6S)-4,5,6-trihydroxy-3-(hydroxymethyl)cyclohex-2-en-1-yl]amino}-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain D:

● Molecule 2: 4,6-dideoxy-4-{[(1S,4R,5S,6S)-4,5,6-trihydroxy-3-(hydroxymethyl)cyclohex-2-en-1-yl]amino}-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

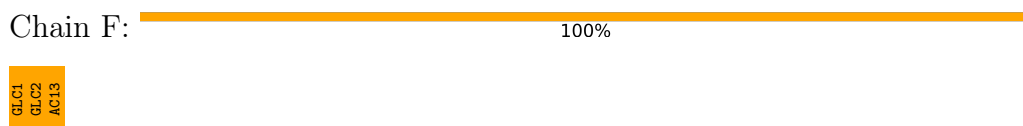Chain E:

● Molecule 2: 4,6-dideoxy-4-{[(1S,4R,5S,6S)-4,5,6-trihydroxy-3-(hydroxymethyl)cyclohex-2-en-1-yl]amino}-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain F:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	116.76Å 163.99Å 311.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.38 – 2.79 58.38 – 2.79	Depositor EDS
% Data completeness (in resolution range)	99.6 (58.38-2.79) 99.6 (58.38-2.79)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.10 (at 2.77Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.194 , 0.239 0.196 , 0.240	Depositor DCC
R_{free} test set	3741 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	48.9	Xtriage
Anisotropy	0.070	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 36.6	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	15585	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.66% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: AC1, GLC

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	0.54	1/5346 (0.0%)	0.64	1/7253 (0.0%)
1	B	0.43	0/5219	0.59	0/7083
1	C	0.44	1/5233 (0.0%)	0.60	0/7102
All	All	0.47	2/15798 (0.0%)	0.61	1/21438 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	473	CYS	CB-SG	-7.45	1.69	1.82
1	C	473	CYS	CB-SG	-5.04	1.73	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	262	ARG	CG-CD-NE	-5.08	101.14	111.80

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	5184	0	4800	85	0
1	B	5064	0	4635	115	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	5076	0	4656	101	0
2	D	44	43	30	1	0
2	E	44	43	30	1	0
2	F	44	43	30	3	0
All	All	15456	129	14181	306	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 10.

The worst 5 of 306 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:306:THR:HG22	1:A:308:SER:H	1.36	0.90
1:B:306:THR:HG22	1:B:308:SER:H	1.37	0.89
1:A:421:CYS:HB2	1:A:471:GLU:HG2	1.60	0.83
1:C:306:THR:HG22	1:C:308:SER:H	1.46	0.81
1:A:143:LYS:HD3	1:A:155:TYR:HB3	1.63	0.80

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	642/668 (96%)	628 (98%)	13 (2%)	1 (0%)	47	78
1	B	631/668 (94%)	618 (98%)	13 (2%)	0	100	100
1	C	633/668 (95%)	615 (97%)	16 (2%)	2 (0%)	41	72
All	All	1906/2004 (95%)	1861 (98%)	42 (2%)	3 (0%)	47	78

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	539	MET
1	C	262	ARG
1	C	555	GLY

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	525/575 (91%)	519 (99%)	6 (1%)	73	92
1	B	504/575 (88%)	497 (99%)	7 (1%)	67	90
1	C	505/575 (88%)	500 (99%)	5 (1%)	76	93
All	All	1534/1725 (89%)	1516 (99%)	18 (1%)	71	92

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	249	PHE
1	C	601	GLN
1	C	578	LYS
1	B	249	PHE
1	C	189	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	601	GLN
1	B	502	ASN
1	C	601	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 ⓘ

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	GLC	D	1	2	12,12,12	0.79	0	17,17,17	2.00	6 (35%)
2	GLC	D	2	2	11,11,12	0.27	0	15,15,17	1.95	1 (6%)
2	AC1	D	3	2	21,22,23	0.89	1 (4%)	22,32,34	1.28	2 (9%)
2	GLC	E	1	2	12,12,12	0.68	0	17,17,17	1.64	2 (11%)
2	GLC	E	2	2	11,11,12	0.49	0	15,15,17	1.82	1 (6%)
2	AC1	E	3	2	21,22,23	0.79	1 (4%)	22,32,34	1.07	2 (9%)
2	GLC	F	1	2	12,12,12	0.79	0	17,17,17	2.19	5 (29%)
2	GLC	F	2	2	11,11,12	0.55	0	15,15,17	1.71	3 (20%)
2	AC1	F	3	2	21,22,23	0.66	0	22,32,34	1.66	4 (18%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GLC	D	1	2	-	2/2/22/22	0/1/1/1
2	GLC	D	2	2	-	0/2/19/22	0/1/1/1
2	AC1	D	3	2	-	4/6/43/46	0/2/2/2
2	GLC	E	1	2	-	2/2/22/22	0/1/1/1
2	GLC	E	2	2	-	0/2/19/22	0/1/1/1
2	AC1	E	3	2	-	3/6/43/46	0/2/2/2
2	GLC	F	1	2	-	1/2/22/22	0/1/1/1
2	GLC	F	2	2	-	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	AC1	F	3	2	-	4/6/43/46	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	3	AC1	C7B-C5B	3.01	1.37	1.32
2	E	3	AC1	C7B-C5B	2.17	1.35	1.32

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	2	GLC	C1-O5-C5	6.53	121.04	112.19
2	E	2	GLC	C1-O5-C5	6.39	120.85	112.19
2	F	2	GLC	C1-O5-C5	4.92	118.86	112.19
2	F	1	GLC	C3-C4-C5	-4.88	101.53	110.24
2	F	1	GLC	O4-C4-C3	4.56	120.90	110.35

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	3	AC1	C4A-C5B-C6B-O6B
2	D	3	AC1	C7B-C5B-C6B-O6B
2	E	3	AC1	C7B-C5B-C6B-O6B
2	F	3	AC1	C7B-C5B-C6B-O6B
2	D	1	GLC	O5-C5-C6-O6

There are no ring outliers.

7 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	1	GLC	1	0
2	D	2	GLC	1	0
2	F	1	GLC	2	0
2	F	3	AC1	1	0
2	E	2	GLC	1	0
2	F	2	GLC	2	0
2	D	1	GLC	1	0

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

Oligosaccharide Chain D

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain F
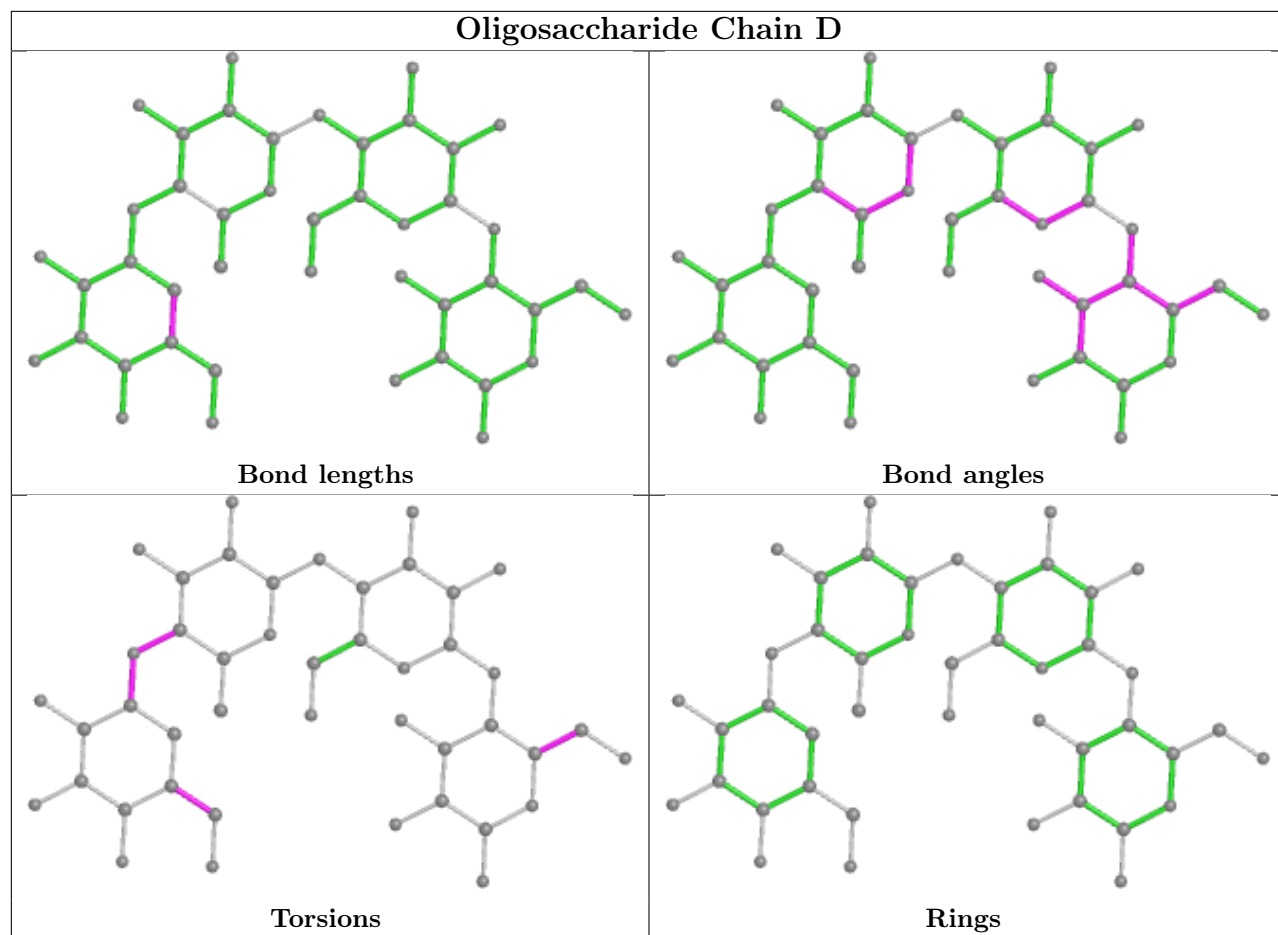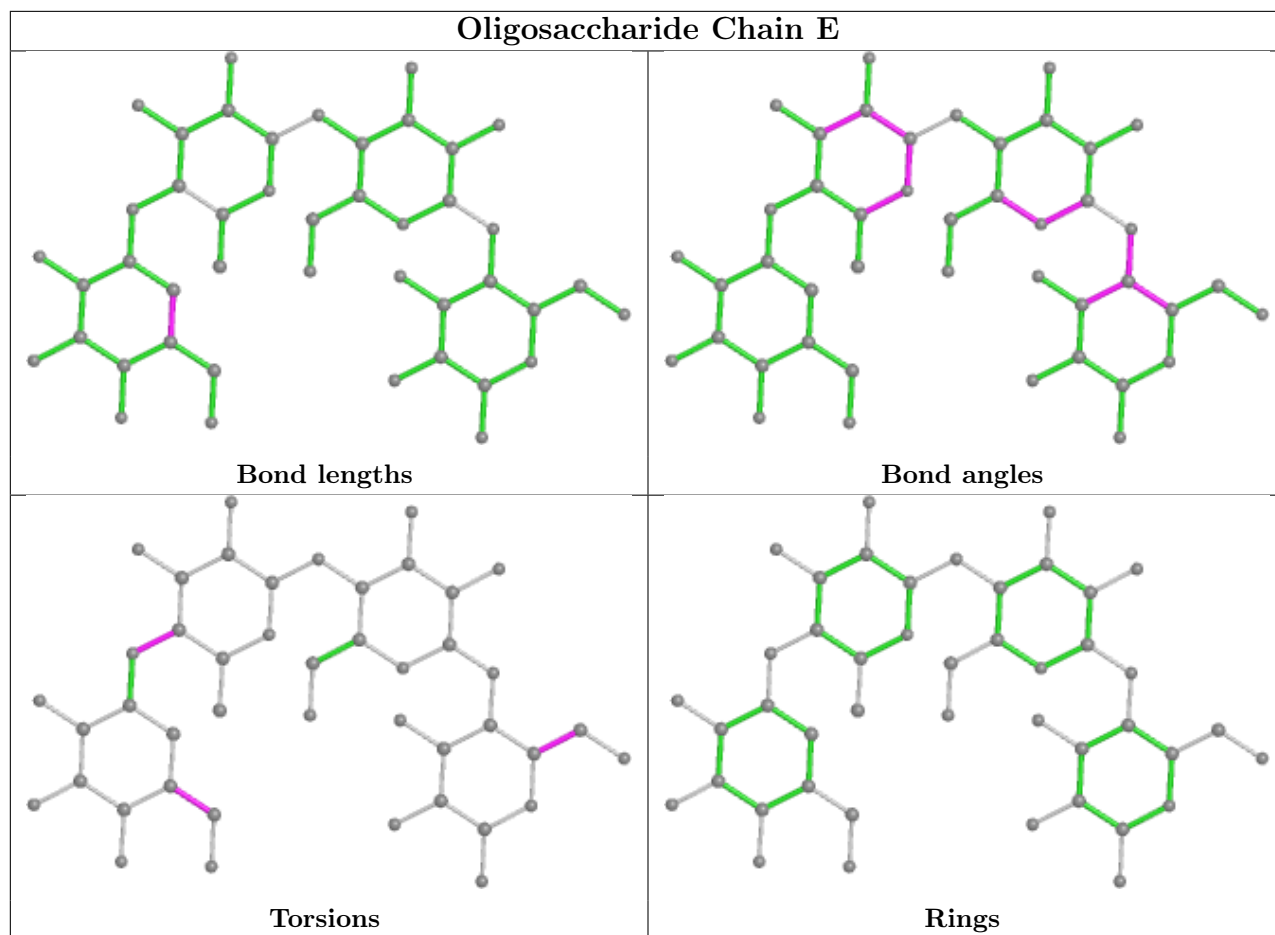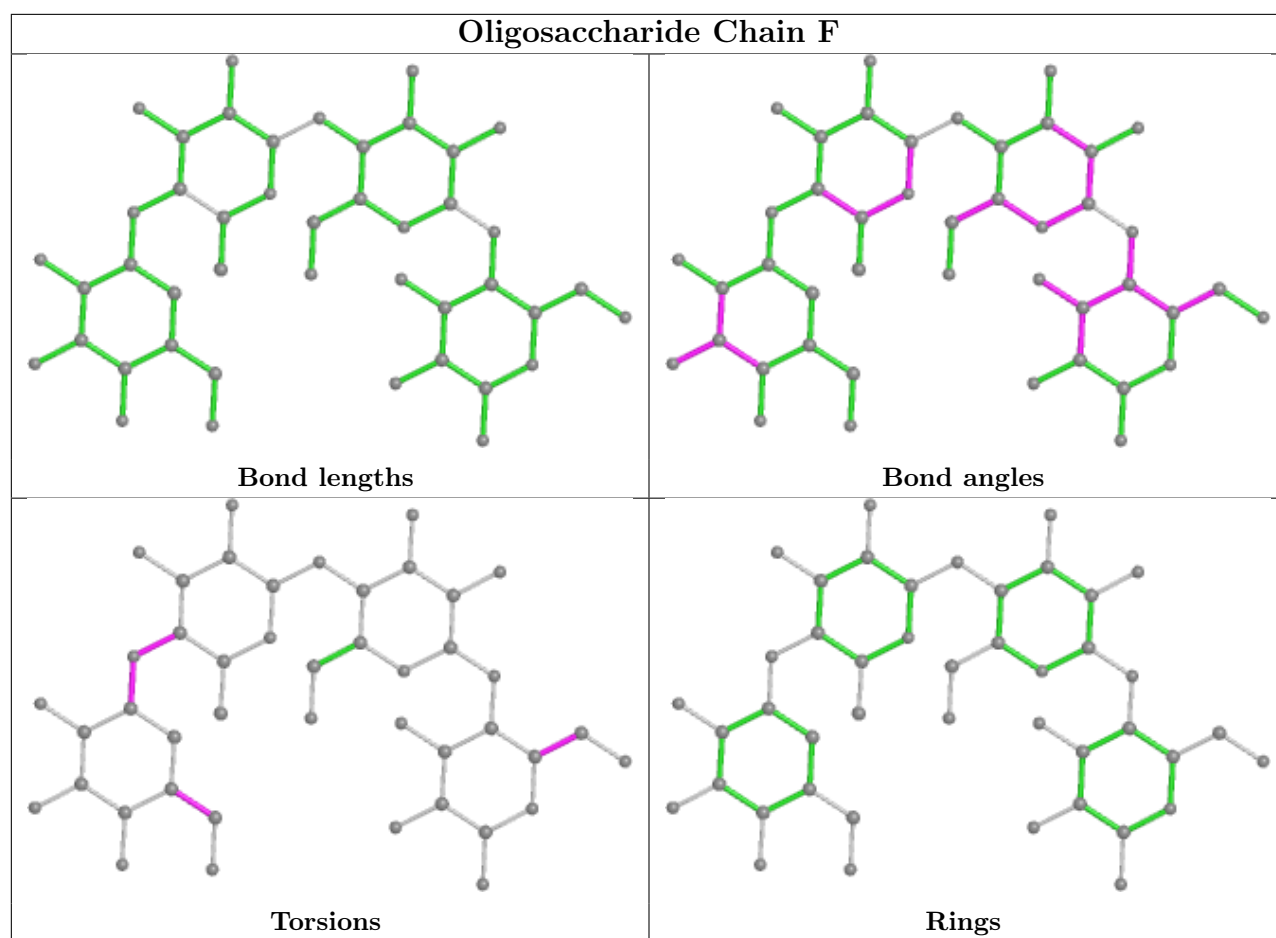
5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

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 ⓘ

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	646/668 (96%)	-0.47	0	100	100	13, 23, 43, 64	0
1	B	637/668 (95%)	-0.28	1 (0%)	95	94	29, 43, 64, 85	0
1	C	637/668 (95%)	-0.16	5 (0%)	86	81	23, 44, 70, 90	0
All	All	1920/2004 (95%)	-0.30	6 (0%)	94	93	13, 38, 65, 90	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	511	PRO	2.5
1	C	635	ASP	2.3
1	C	509	PHE	2.2
1	B	150	SER	2.2
1	C	579	PHE	2.2

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

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

6.3 Carbohydrates ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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.
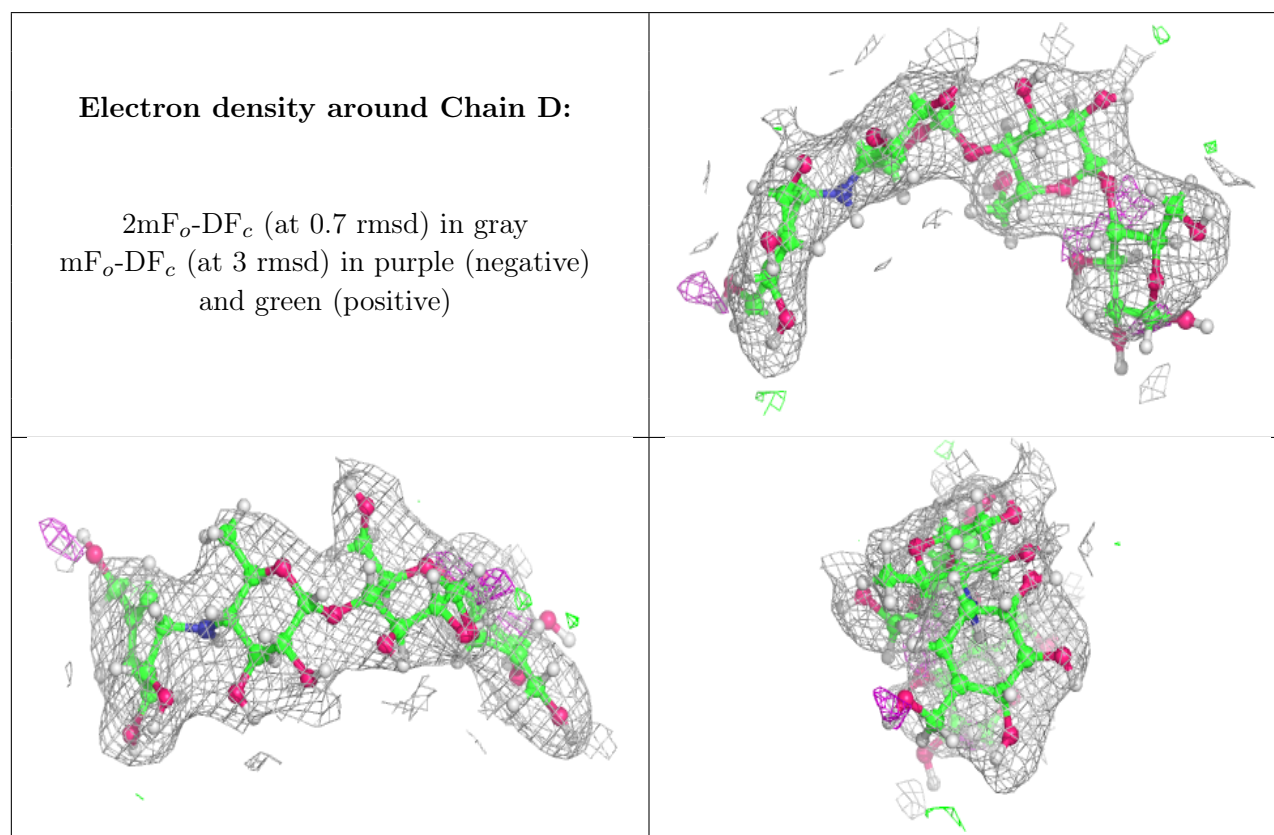
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GLC	F	1	12/12	0.79	0.40	66,81,104,121	0
2	GLC	E	1	12/12	0.85	0.40	55,69,90,100	0
2	GLC	D	1	12/12	0.90	0.21	25,47,76,91	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
2	AC1	F	3	21/22	0.90	0.21	47,60,81,97	0
2	GLC	F	2	11/12	0.92	0.20	52,64,77,81	0
2	AC1	E	3	21/22	0.94	0.19	36,53,85,97	0
2	GLC	E	2	11/12	0.94	0.20	42,51,61,66	0
2	AC1	D	3	21/22	0.96	0.16	20,27,50,63	0
2	GLC	D	2	11/12	0.97	0.13	23,29,36,41	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.
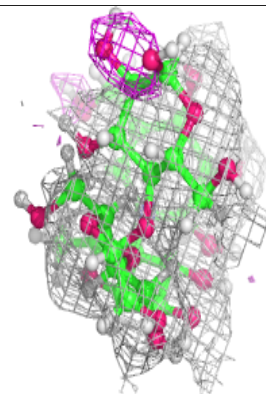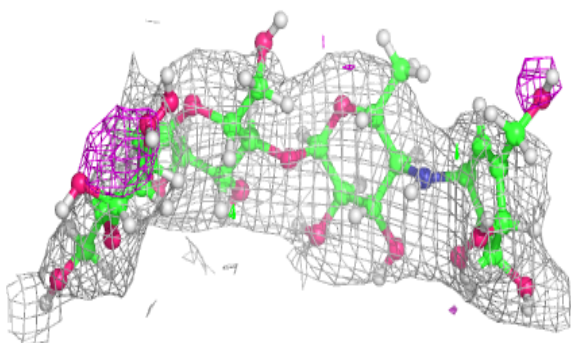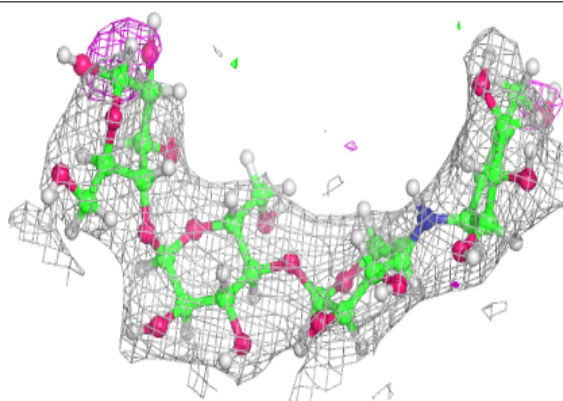

Electron density around Chain D:

$2mF_o$-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 E:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
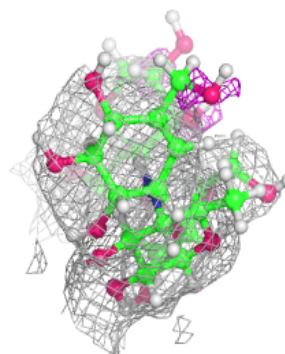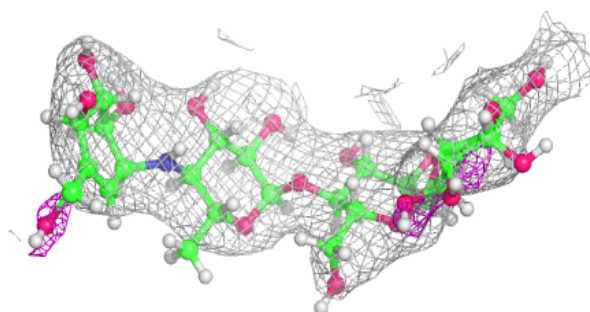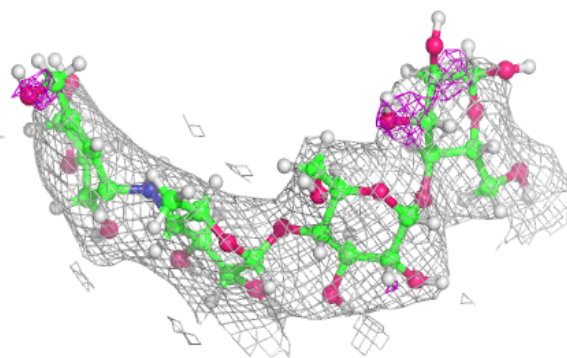
Electron density around Chain F:

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 ⓘ

There are no ligands in this entry.

6.5   Other polymers ⓘ

There are no such residues in this entry.