



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 06:15 PM UTC

PDB ID : 4FNT / pdb_00004fnt
Title : Crystal structure of GH36 alpha-galactosidase AgaA A355E D548N from Geobacillus stearothermophilus in complex with raffinose
Authors : Merceron, R.; Foucault, M.; Haser, R.; Mattes, R.; Watzlawick, H.; Gouet, P.
Deposited on : 2012-06-20
Resolution : 2.60 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

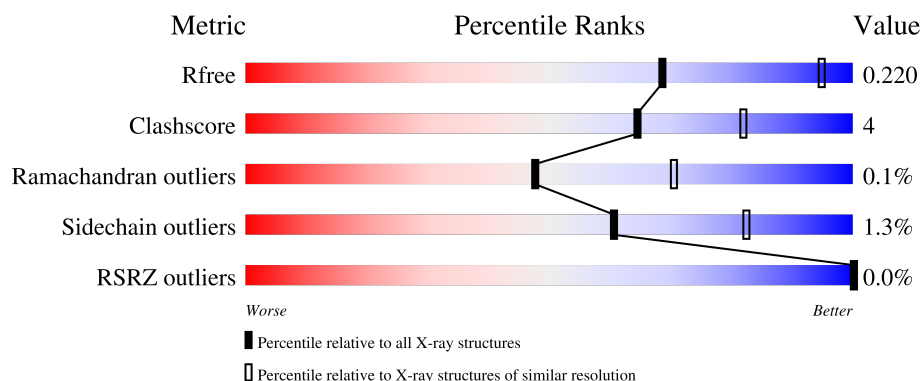
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	4008 (2.60-2.60)
Clashscore	190562	4347 (2.60-2.60)
Ramachandran outliers	187476	4277 (2.60-2.60)
Sidechain outliers	187428	4277 (2.60-2.60)
RSRZ outliers	180081	4008 (2.60-2.60)

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	729	87% 10% ..
1	B	729	88% 9% ..
1	C	729	89% 9% ..
1	D	729	89% 9% .
2	E	3	67% 33%

Continued on next page...

Continued from previous page...

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

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 23619 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 Alpha-galactosidase AgaA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	718	Total	C	N	O	S	0	0	0
			5790	3676	1021	1071	22			
1	B	718	Total	C	N	O	S	0	0	0
			5790	3676	1021	1071	22			
1	C	718	Total	C	N	O	S	0	0	0
			5790	3676	1021	1071	22			
1	D	718	Total	C	N	O	S	0	0	0
			5790	3676	1021	1071	22			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	355	GLU	ALA	engineered mutation	UNP Q9ALJ4
A	548	ASN	ASP	engineered mutation	UNP Q9ALJ4
A	704	VAL	MET	engineered mutation	UNP Q9ALJ4
B	355	GLU	ALA	engineered mutation	UNP Q9ALJ4
B	548	ASN	ASP	engineered mutation	UNP Q9ALJ4
B	704	VAL	MET	engineered mutation	UNP Q9ALJ4
C	355	GLU	ALA	engineered mutation	UNP Q9ALJ4
C	548	ASN	ASP	engineered mutation	UNP Q9ALJ4
C	704	VAL	MET	engineered mutation	UNP Q9ALJ4
D	355	GLU	ALA	engineered mutation	UNP Q9ALJ4
D	548	ASN	ASP	engineered mutation	UNP Q9ALJ4
D	704	VAL	MET	engineered mutation	UNP Q9ALJ4

- Molecule 2 is an oligosaccharide called alpha-D-galactopyranose-(1-6)-alpha-D-glucopyranoside-(1-2)-beta-D-fructofuranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	E	3	Total	C	O	0	0	0
			34	18	16			
2	F	3	Total	C	O	0	0	0
			34	18	16			
2	G	3	Total	C	O	0	0	0
			34	18	16			

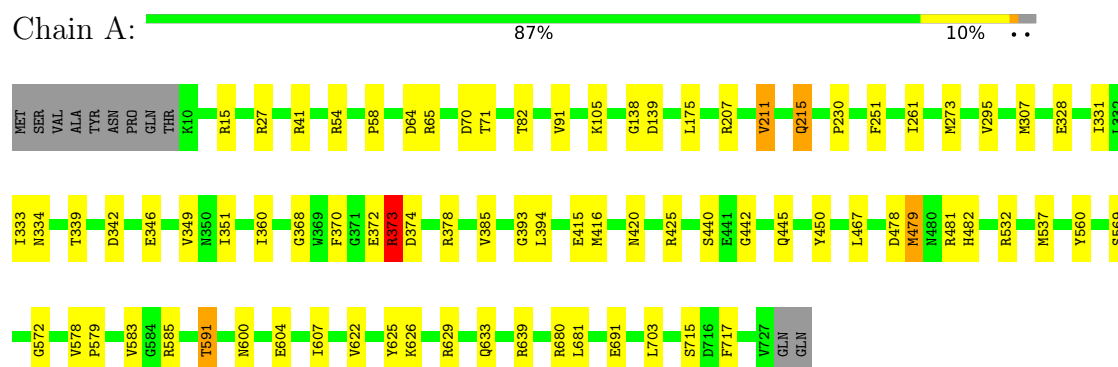
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	114	Total	O	0	0
			114	114		
3	B	86	Total	O	0	0
			86	86		
3	C	49	Total	O	0	0
			49	49		
3	D	108	Total	O	0	0
			108	108		

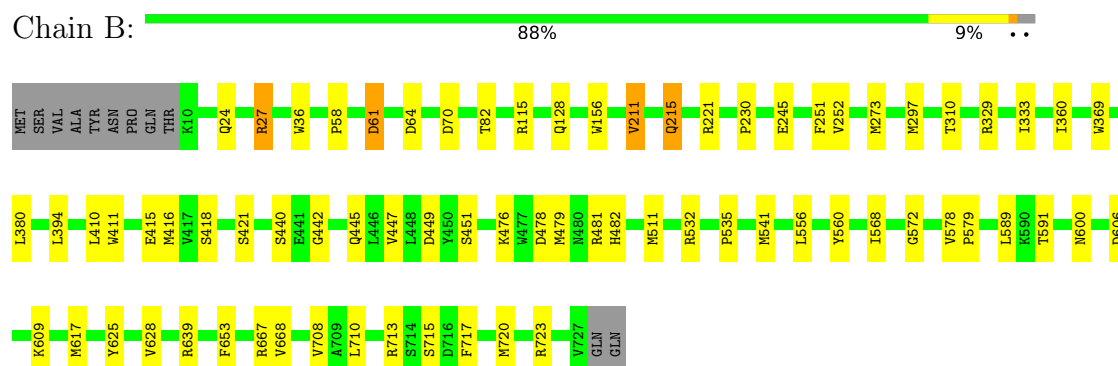
3 Residue-property plots [i](#)

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.

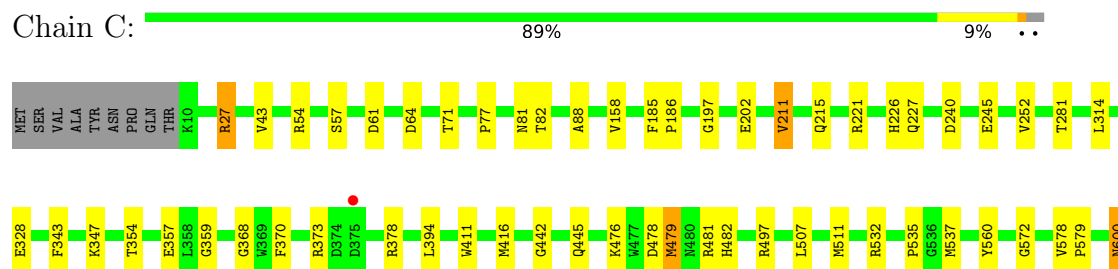
• Molecule 1: Alpha-galactosidase AgaA



• Molecule 1: Alpha-galactosidase AgaA



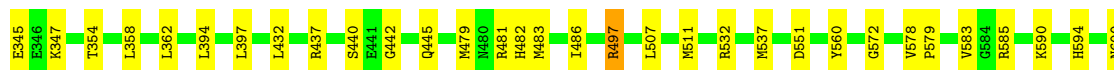
• Molecule 1: Alpha-galactosidase AgaA





- Molecule 1: Alpha-galactosidase AgaA

Chain D: 89% 9%



- Molecule 2: alpha-D-galactopyranose-(1-6)-alpha-D-glucopyranose-(1-2)-beta-D-fructofuranose

Chain E: 67% 33%



- Molecule 2: alpha-D-galactopyranose-(1-6)-alpha-D-glucopyranose-(1-2)-beta-D-fructofuranose

Chain F: 67% 33%



- Molecule 2: alpha-D-galactopyranose-(1-6)-alpha-D-glucopyranose-(1-2)-beta-D-fructofuranose

Chain G: 100%



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	149.68Å 149.68Å 234.77Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.89 – 2.60 19.89 – 2.60	Depositor EDS
% Data completeness (in resolution range)	98.8 (19.89-2.60) 98.5 (19.89-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.09 (at 2.59Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.170 , 0.221 0.170 , 0.220	Depositor DCC
R_{free} test set	4629 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	46.9	Xtriage
Anisotropy	0.161	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 33.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.014 for -h,-k,l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	23619	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.38% 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: FRU, GLA, 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	0.29	0/5937	0.70	0/8050
1	B	0.27	0/5937	0.70	3/8050 (0.0%)
1	C	0.27	0/5937	0.71	5/8050 (0.1%)
1	D	0.29	0/5937	0.70	0/8050
All	All	0.28	0/23748	0.70	8/32200 (0.0%)

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	61	ASP	CA-C-N	6.46	126.96	119.47
1	B	61	ASP	C-N-CA	6.46	126.96	119.47
1	C	27	ARG	CB-CA-C	-5.55	110.19	116.63
1	B	27	ARG	CB-CA-C	-5.47	110.25	116.54
1	C	61	ASP	CA-C-N	5.21	124.87	119.56

There are no chirality outliers.

There are no planarity outliers.

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	5790	0	5622	50	0
1	B	5790	0	5622	43	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	5790	0	5622	42	0
1	D	5790	0	5622	44	0
2	E	34	0	30	1	0
2	F	34	0	30	1	0
2	G	34	0	27	5	0
3	A	114	0	0	0	0
3	B	86	0	0	1	0
3	C	49	0	0	2	0
3	D	108	0	0	0	0
All	All	23619	0	22575	177	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 4.

The worst 5 of 177 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:D:511:MET:HE1	1:D:537:MET:HE1	1.65	0.79
1:D:328:GLU:HB3	1:D:629:ARG:HD3	1.67	0.76
1:D:481:ARG:NH1	1:D:482:HIS:O	2.21	0.72
1:C:481:ARG:NH1	1:C:482:HIS:O	2.22	0.72
1:A:440:SER:O	1:A:481:ARG:NH2	2.22	0.72

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	716/729 (98%)	691 (96%)	23 (3%)	2 (0%)	36 58
1	B	716/729 (98%)	689 (96%)	27 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	716/729 (98%)	690 (96%)	26 (4%)	0	100	100
1	D	716/729 (98%)	686 (96%)	30 (4%)	0	100	100
All	All	2864/2916 (98%)	2756 (96%)	106 (4%)	2 (0%)	48	70

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	373	ARG
1	A	715	SER

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	609/619 (98%)	600 (98%)	9 (2%)	57	80
1	B	609/619 (98%)	599 (98%)	10 (2%)	55	79
1	C	609/619 (98%)	603 (99%)	6 (1%)	68	86
1	D	609/619 (98%)	603 (99%)	6 (1%)	68	86
All	All	2436/2476 (98%)	2405 (99%)	31 (1%)	61	82

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	394	LEU
1	D	394	LEU
1	B	708	VAL
1	D	583	VAL
1	C	613	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	392	ASN
1	D	260	GLN
1	D	688	GLN
1	D	480	ASN
1	A	334	ASN

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	FRU	E	1	2	11,12,12	1.66	5 (45%)	10,18,18	1.52	2 (20%)
2	GLC	E	2	2	11,11,12	1.73	3 (27%)	15,15,17	2.02	5 (33%)
2	GLA	E	3	2	11,11,12	1.76	2 (18%)	15,15,17	1.46	4 (26%)
2	FRU	F	1	2	11,12,12	1.70	4 (36%)	10,18,18	1.57	3 (30%)
2	GLC	F	2	2	11,11,12	1.94	3 (27%)	15,15,17	2.55	7 (46%)
2	GLA	F	3	2	11,11,12	1.61	2 (18%)	15,15,17	1.16	2 (13%)
2	FRU	G	1	2	11,12,12	1.70	3 (27%)	10,18,18	1.23	0
2	GLC	G	2	2	11,11,12	2.23	5 (45%)	15,15,17	2.50	8 (53%)
2	GLA	G	3	2	11,11,12	2.69	3 (27%)	15,15,17	1.50	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	FRU	E	1	2	-	0/5/24/24	0/1/1/1
2	GLC	E	2	2	-	1/2/19/22	0/1/1/1
2	GLA	E	3	2	-	0/2/19/22	0/1/1/1
2	FRU	F	1	2	-	3/5/24/24	0/1/1/1
2	GLC	F	2	2	-	2/2/19/22	0/1/1/1
2	GLA	F	3	2	-	0/2/19/22	0/1/1/1
2	FRU	G	1	2	-	5/5/24/24	0/1/1/1
2	GLC	G	2	2	-	2/2/19/22	0/1/1/1
2	GLA	G	3	2	-	1/2/19/22	0/1/1/1

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	3	GLA	C2-C3	-6.17	1.43	1.52
2	G	2	GLC	C2-C3	-4.78	1.45	1.52
2	E	3	GLA	C2-C3	-4.69	1.45	1.52
2	G	3	GLA	C4-C3	-4.64	1.40	1.52
2	F	2	GLC	O5-C1	4.53	1.51	1.43

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	2	GLC	C1-O5-C5	5.27	119.24	112.19
2	F	2	GLC	C1-O5-C5	5.12	119.04	112.19
2	F	2	GLC	C1-C2-C3	4.85	116.70	109.64
2	E	2	GLC	C3-C4-C5	4.83	119.00	110.23
2	G	2	GLC	C3-C4-C5	3.76	117.04	110.23

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

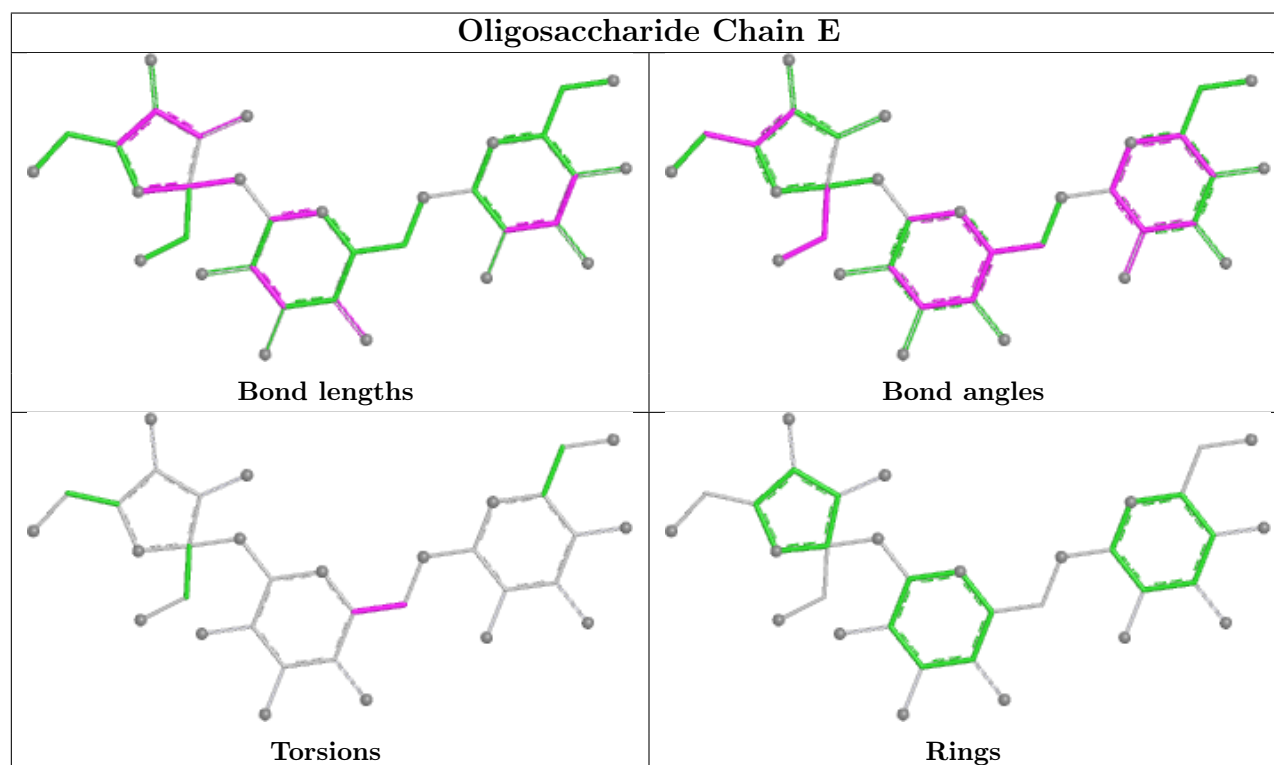
Mol	Chain	Res	Type	Atoms
2	G	1	FRU	O1-C1-C2-C3
2	G	1	FRU	O1-C1-C2-O2
2	G	1	FRU	O1-C1-C2-O5
2	G	1	FRU	C4-C5-C6-O6
2	G	1	FRU	O5-C5-C6-O6

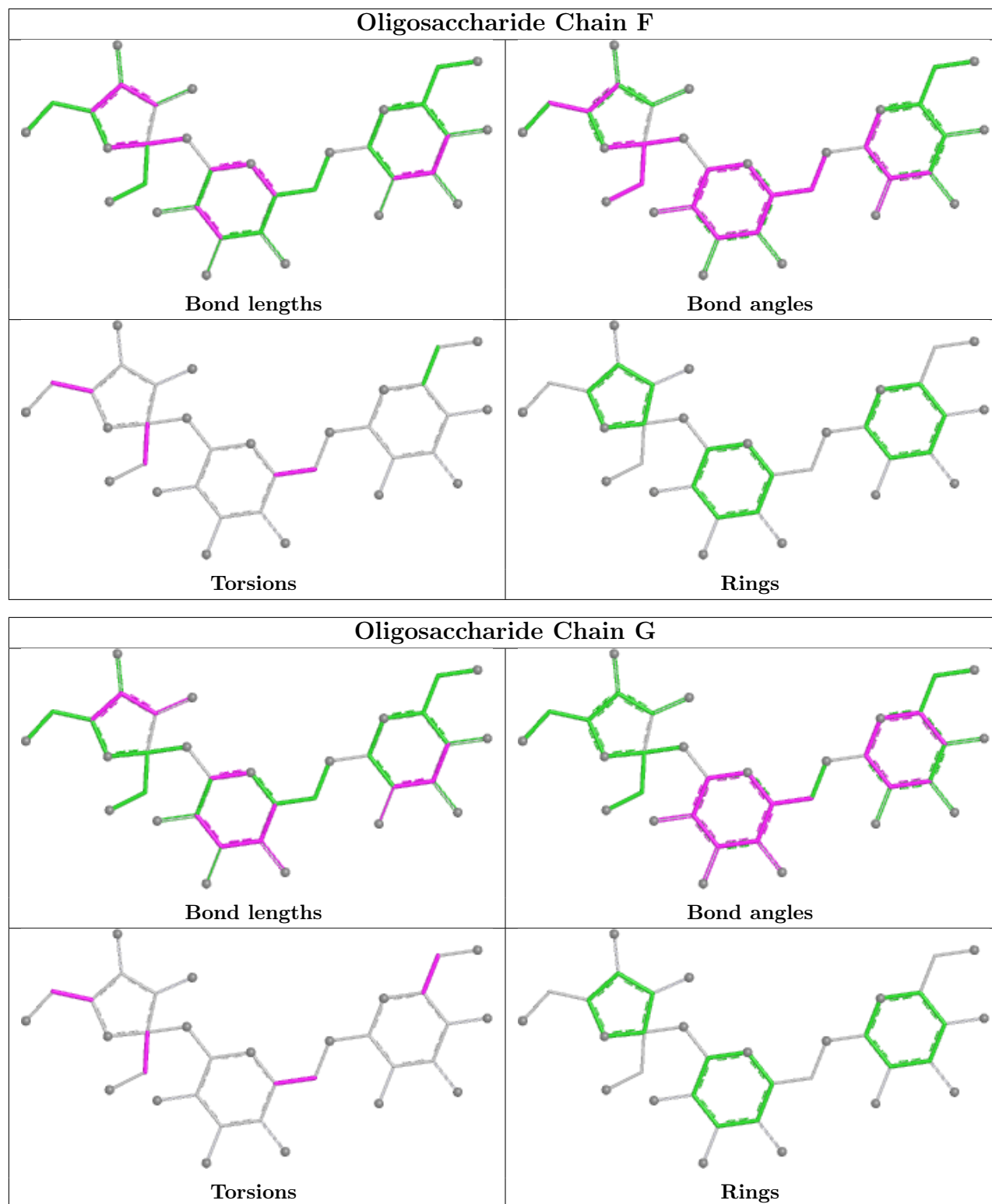
There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	3	GLA	1	0
2	G	3	GLA	1	0
2	G	1	FRU	4	0
2	G	2	GLC	3	0
2	F	3	GLA	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 ⓘ

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	718/729 (98%)	-0.76	0 100 100	26, 37, 59, 88	0
1	B	718/729 (98%)	-0.71	0 100 100	27, 41, 64, 84	0
1	C	718/729 (98%)	-0.57	1 (0%) 92 90	29, 48, 68, 86	0
1	D	718/729 (98%)	-0.82	0 100 100	26, 36, 53, 66	0
All	All	2872/2916 (98%)	-0.71	1 (0%) 100 100	26, 40, 62, 88	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	375	ASP	2.4

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.

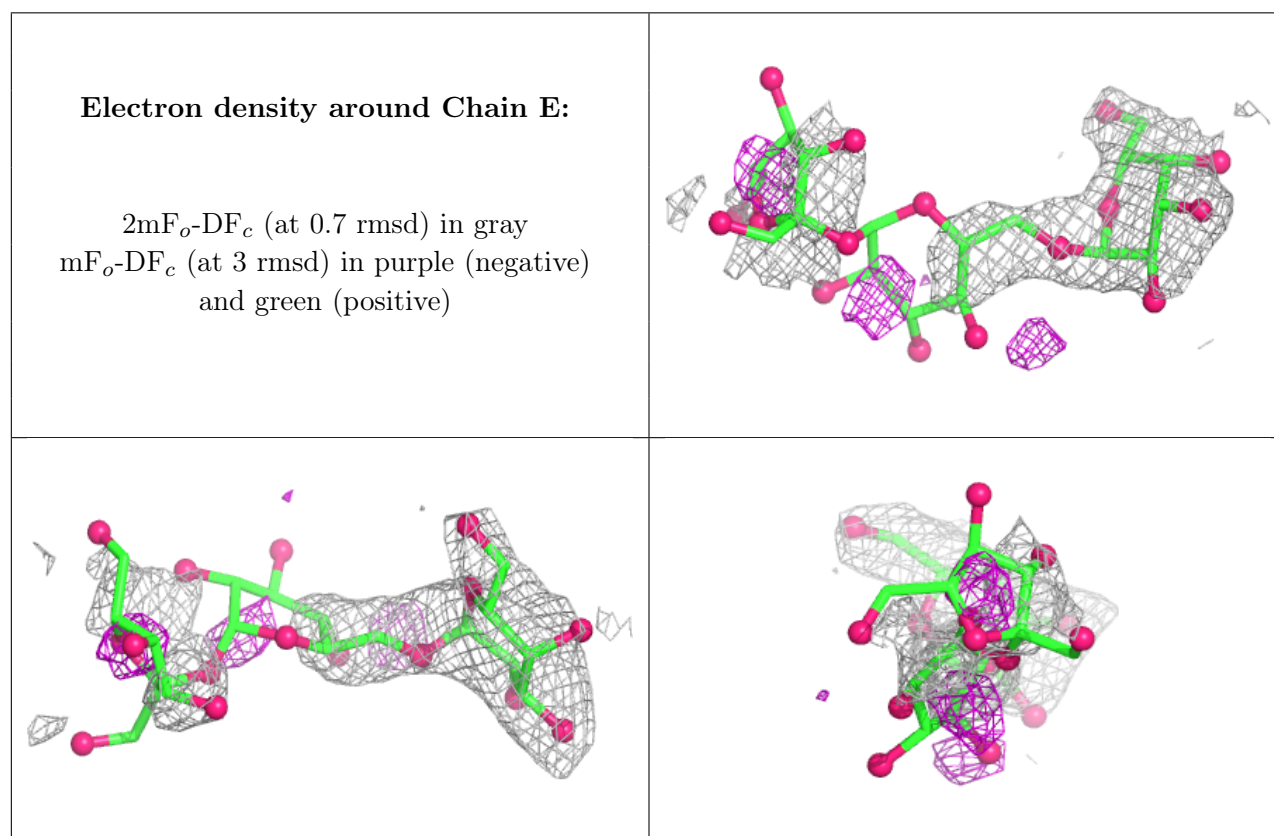
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	FRU	E	1	12/12	0.69	0.14	95,117,128,131	0
2	FRU	G	1	12/12	0.73	0.15	80,101,108,120	0
2	FRU	F	1	12/12	0.76	0.13	80,111,122,129	0
2	GLC	E	2	11/12	0.88	0.16	71,98,109,118	0
2	GLC	G	2	11/12	0.88	0.12	56,79,95,100	0
2	GLA	G	3	11/12	0.90	0.08	44,46,54,55	0

Continued on next page...

Continued from previous page...

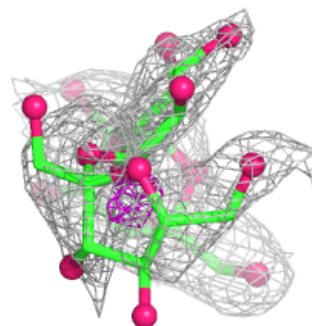
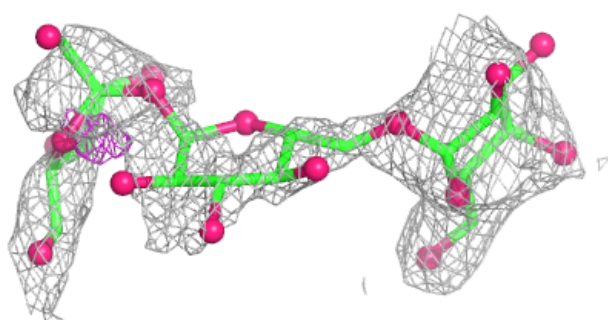
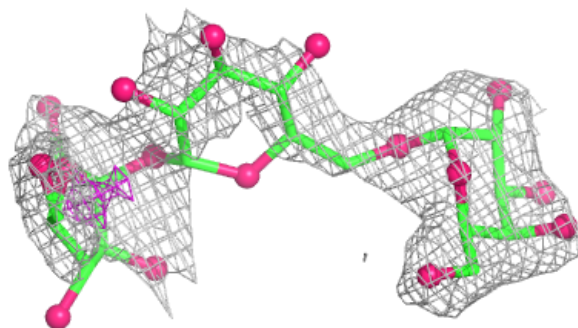
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	GLA	E	3	11/12	0.92	0.08	54,66,69,77	0
2	GLC	F	2	11/12	0.92	0.09	75,87,98,105	0
2	GLA	F	3	11/12	0.92	0.07	46,61,71,73	0

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.

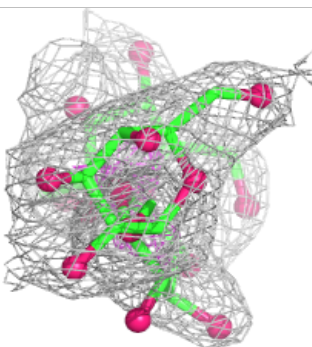
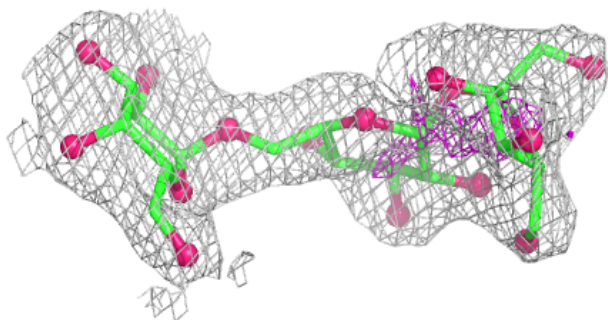
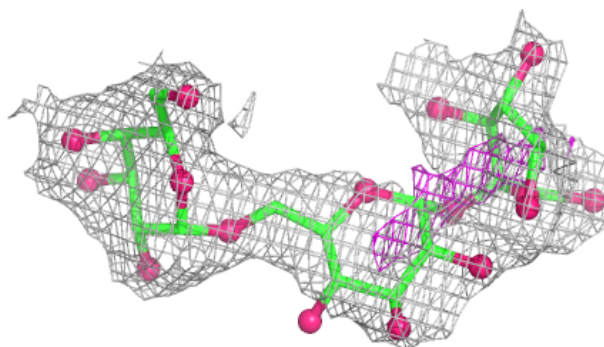


Electron density around Chain F:

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

**Electron density around Chain G:**

$2mF_o-DF_c$ (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.