



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 06:17 PM UTC

PDB ID : 3QHN / pdb_00003qhn
Title : Crystal analysis of the complex structure, E201A-cellobetraose, of endocellulase from pyrococcus horikoshii
Authors : Kim, H.-W.; Ishikawa, K.
Deposited on : 2011-01-26
Resolution : 1.99 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

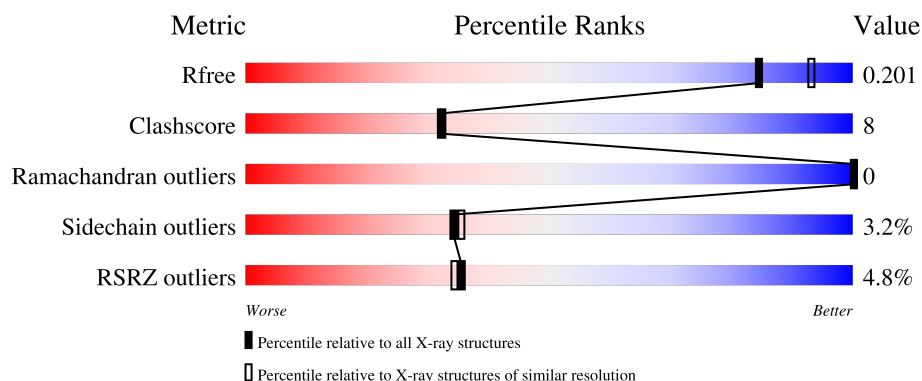
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.99 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	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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	458	
1	B	458	
1	C	458	
2	D	4	
2	E	4	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	4	 50%50%
3	F	2	 50%50%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9624 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 458aa long hypothetical endo-1,4-beta-glucanase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	377	Total	C	N	O	S	0	0	0
			3067	2000	502	555	10			
1	B	377	Total	C	N	O	S	0	0	0
			3067	2000	502	555	10			
1	C	377	Total	C	N	O	S	0	0	0
			3067	2000	502	555	10			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	201	ALA	GLU	engineered mutation	UNP O58925
B	201	ALA	GLU	engineered mutation	UNP O58925
C	201	ALA	GLU	engineered mutation	UNP O58925

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	D	4	Total	C	O	0	0	0
			45	24	21			
2	E	4	Total	C	O	0	0	0
			45	24	21			
2	G	4	Total	C	O	0	0	0
			45	24	21			

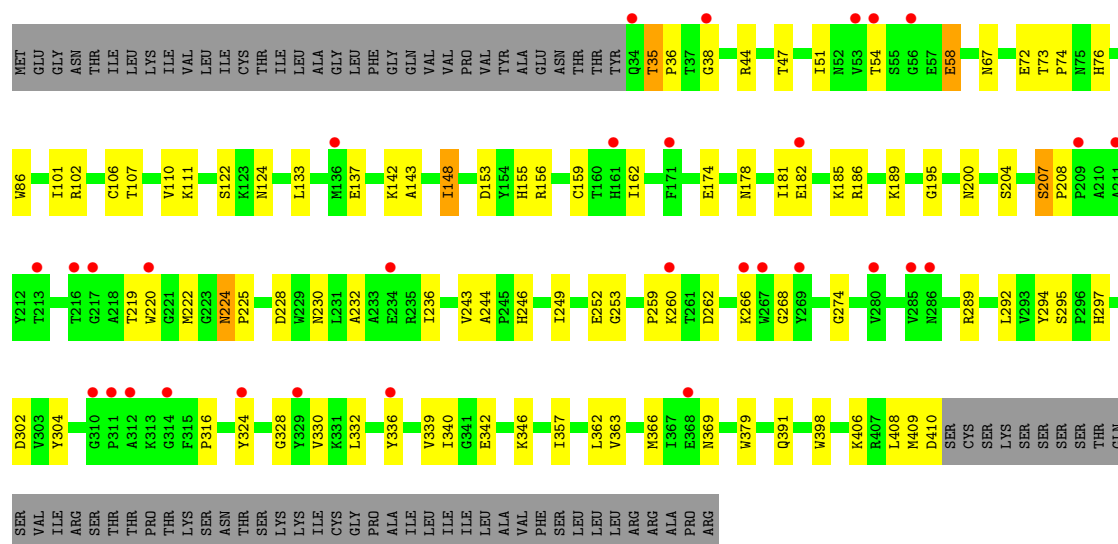
- Molecule 3 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose.



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	108	Total	O	0	0
			108	108		
4	B	106	Total	O	0	0
			106	106		
4	C	51	Total	O	0	0
			51	51		



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

Chain D: 75% 25%



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

Chain E: 50% 50%



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

Chain G: 50% 50%



- Molecule 3: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain F: 50% 50%



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	162.32Å 58.35Å 138.27Å 90.00° 109.61° 90.00°	Depositor
Resolution (Å)	28.99 – 1.99 28.99 – 1.99	Depositor EDS
% Data completeness (in resolution range)	94.8 (28.99-1.99) 90.7 (28.99-1.99)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.01 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.196 , 0.248 0.198 , 0.201	Depositor DCC
R_{free} test set	3988 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	23.3	Xtriage
Anisotropy	0.436	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 45.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	9624	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.55% 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: 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	1.26	7/3178 (0.2%)	1.12	2/4339 (0.0%)
1	B	1.34	8/3178 (0.3%)	1.10	7/4339 (0.2%)
1	C	1.09	4/3178 (0.1%)	1.08	9/4339 (0.2%)
All	All	1.23	19/9534 (0.2%)	1.10	18/13017 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	201	ALA	CA-C	6.97	1.60	1.53
1	A	232	ALA	CA-CB	-6.92	1.42	1.53
1	C	162	ILE	CA-CB	6.53	1.61	1.53
1	B	201	ALA	CA-CB	6.45	1.58	1.52
1	A	219	THR	CA-CB	6.33	1.63	1.53

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	224	ASN	CA-C-N	7.77	129.55	119.84
1	C	224	ASN	C-N-CA	7.77	129.55	119.84
1	C	195	GLY	N-CA-C	7.36	119.81	110.38
1	B	170	ASP	N-CA-C	6.39	120.39	112.59
1	C	244	ALA	CA-C-N	5.98	126.40	119.47

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	3067	0	2907	51	0
1	B	3067	0	2907	33	1
1	C	3067	0	2907	59	0
2	D	45	0	39	1	0
2	E	45	0	39	2	0
2	G	45	0	39	2	1
3	F	23	0	21	2	0
4	A	108	0	0	2	0
4	B	106	0	0	0	0
4	C	51	0	0	0	0
All	All	9624	0	8859	146	1

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 8.

The worst 5 of 146 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:B:316:PRO:HG3	1:B:357:ILE:CD1	1.87	1.04
1:B:316:PRO:HG3	1:B:357:ILE:HD13	1.53	0.89
1:C:266:LYS:HD2	1:C:266:LYS:H	1.34	0.89
1:B:67:ASN:HD21	1:B:102:ARG:HH11	1.24	0.86
1:C:67:ASN:HD21	1:C:102:ARG:HH11	1.29	0.80

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:42:GLU:OE2	2:G:4:BGC:O4[4_545]	2.15	0.05

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	375/458 (82%)	362 (96%)	13 (4%)	0	100	100
1	B	375/458 (82%)	365 (97%)	10 (3%)	0	100	100
1	C	375/458 (82%)	361 (96%)	14 (4%)	0	100	100
All	All	1125/1374 (82%)	1088 (97%)	37 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	323/395 (82%)	314 (97%)	9 (3%)	38	41
1	B	323/395 (82%)	315 (98%)	8 (2%)	42	45
1	C	323/395 (82%)	309 (96%)	14 (4%)	26	25
All	All	969/1185 (82%)	938 (97%)	31 (3%)	34	35

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

Mol	Chain	Res	Type
1	B	313	LYS
1	C	243	VAL
1	C	35	THR
1	C	289	ARG
1	C	189	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	67	ASN
1	C	161	HIS
1	C	359	GLN
1	C	200	ASN
1	C	124	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 ⓘ

14 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	D	1	2	12,12,12	1.03	1 (8%)	17,17,17	1.28	0
2	BGC	D	2	2	11,11,12	1.16	1 (9%)	15,15,17	1.45	3 (20%)
2	BGC	D	3	2	11,11,12	2.59	4 (36%)	15,15,17	2.36	5 (33%)
2	BGC	D	4	2	11,11,12	1.28	1 (9%)	15,15,17	1.21	1 (6%)
2	BGC	E	1	2	12,12,12	1.67	3 (25%)	17,17,17	1.60	3 (17%)
2	BGC	E	2	2	11,11,12	1.30	1 (9%)	15,15,17	1.38	2 (13%)
2	BGC	E	3	2	11,11,12	1.63	3 (27%)	15,15,17	2.30	8 (53%)
2	BGC	E	4	2	11,11,12	0.96	1 (9%)	15,15,17	0.90	1 (6%)
3	BGC	F	1	3	12,12,12	0.54	0	17,17,17	0.84	1 (5%)
3	BGC	F	2	3	11,11,12	0.54	0	15,15,17	0.78	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	BGC	G	1	2	12,12,12	1.45	2 (16%)	17,17,17	2.28	7 (41%)
2	BGC	G	2	2	11,11,12	0.91	0	15,15,17	1.41	2 (13%)
2	BGC	G	3	2	11,11,12	0.99	2 (18%)	15,15,17	1.76	1 (6%)
2	BGC	G	4	2	11,11,12	1.35	2 (18%)	15,15,17	2.12	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
2	BGC	D	1	2	-	0/2/22/22	0/1/1/1
2	BGC	D	2	2	-	1/2/19/22	0/1/1/1
2	BGC	D	3	2	-	0/2/19/22	0/1/1/1
2	BGC	D	4	2	-	0/2/19/22	0/1/1/1
2	BGC	E	1	2	-	2/2/22/22	0/1/1/1
2	BGC	E	2	2	-	0/2/19/22	0/1/1/1
2	BGC	E	3	2	-	2/2/19/22	0/1/1/1
2	BGC	E	4	2	-	1/2/19/22	0/1/1/1
3	BGC	F	1	3	-	0/2/22/22	0/1/1/1
3	BGC	F	2	3	-	0/2/19/22	0/1/1/1
2	BGC	G	1	2	-	2/2/22/22	0/1/1/1
2	BGC	G	2	2	-	0/2/19/22	0/1/1/1
2	BGC	G	3	2	-	1/2/19/22	0/1/1/1
2	BGC	G	4	2	-	0/2/19/22	0/1/1/1

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	3	BGC	O5-C1	6.23	1.54	1.43
2	E	2	BGC	O5-C1	3.89	1.50	1.43
2	E	3	BGC	O5-C1	3.50	1.49	1.43
2	G	1	BGC	O5-C1	3.24	1.50	1.42
2	D	3	BGC	O5-C5	3.05	1.49	1.43

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	3	BGC	C1-O5-C5	6.57	120.99	112.19
2	G	3	BGC	O4-C4-C3	-5.72	96.90	110.38

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	1	BGC	C6-C5-C4	-5.51	99.49	113.02
2	G	4	BGC	C1-O5-C5	-4.90	105.62	112.19
2	E	3	BGC	C3-C4-C5	4.53	118.45	110.23

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

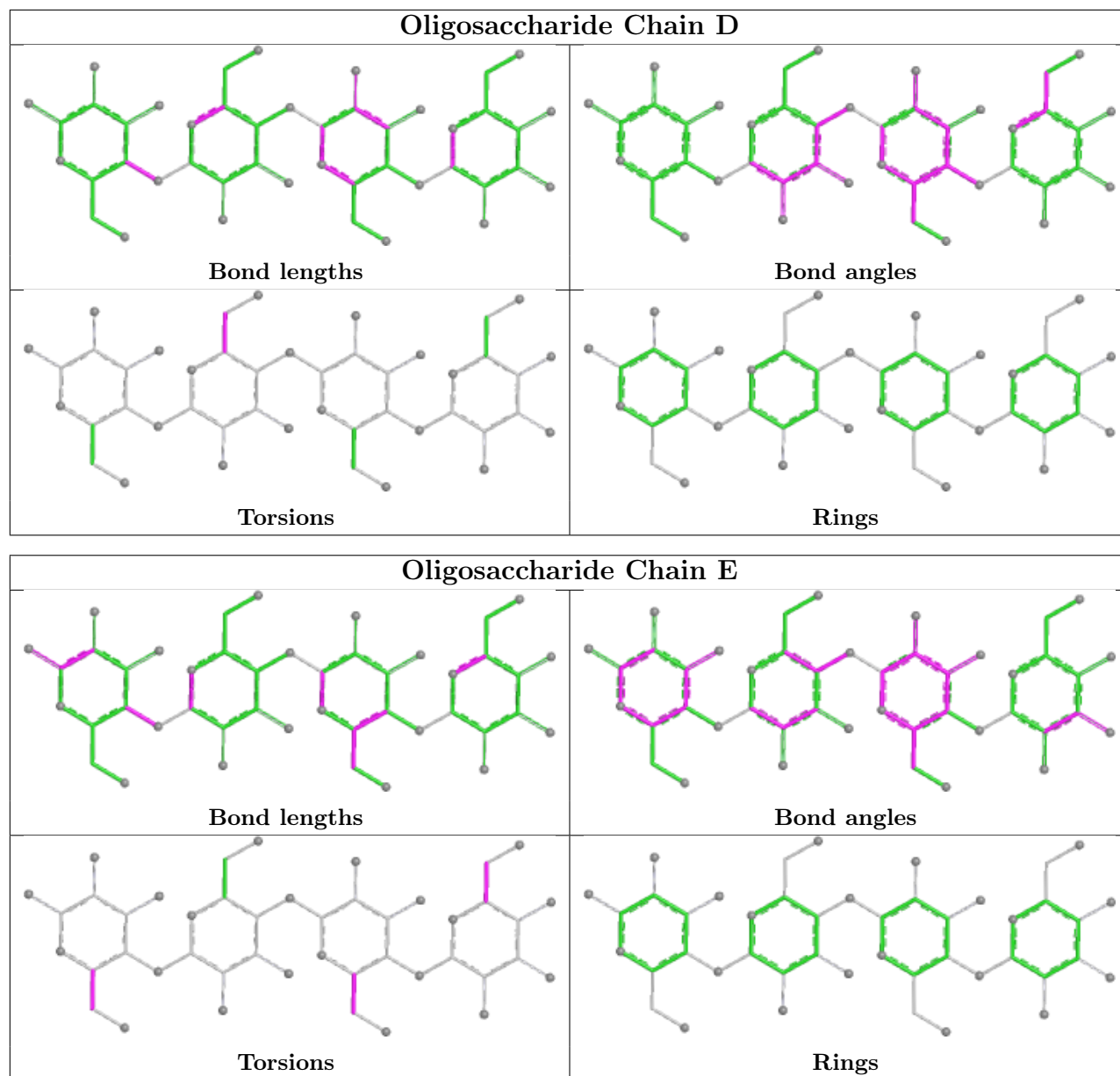
Mol	Chain	Res	Type	Atoms
2	E	1	BGC	O5-C5-C6-O6
2	E	1	BGC	C4-C5-C6-O6
2	G	1	BGC	C4-C5-C6-O6
2	G	1	BGC	O5-C5-C6-O6
2	G	3	BGC	O5-C5-C6-O6

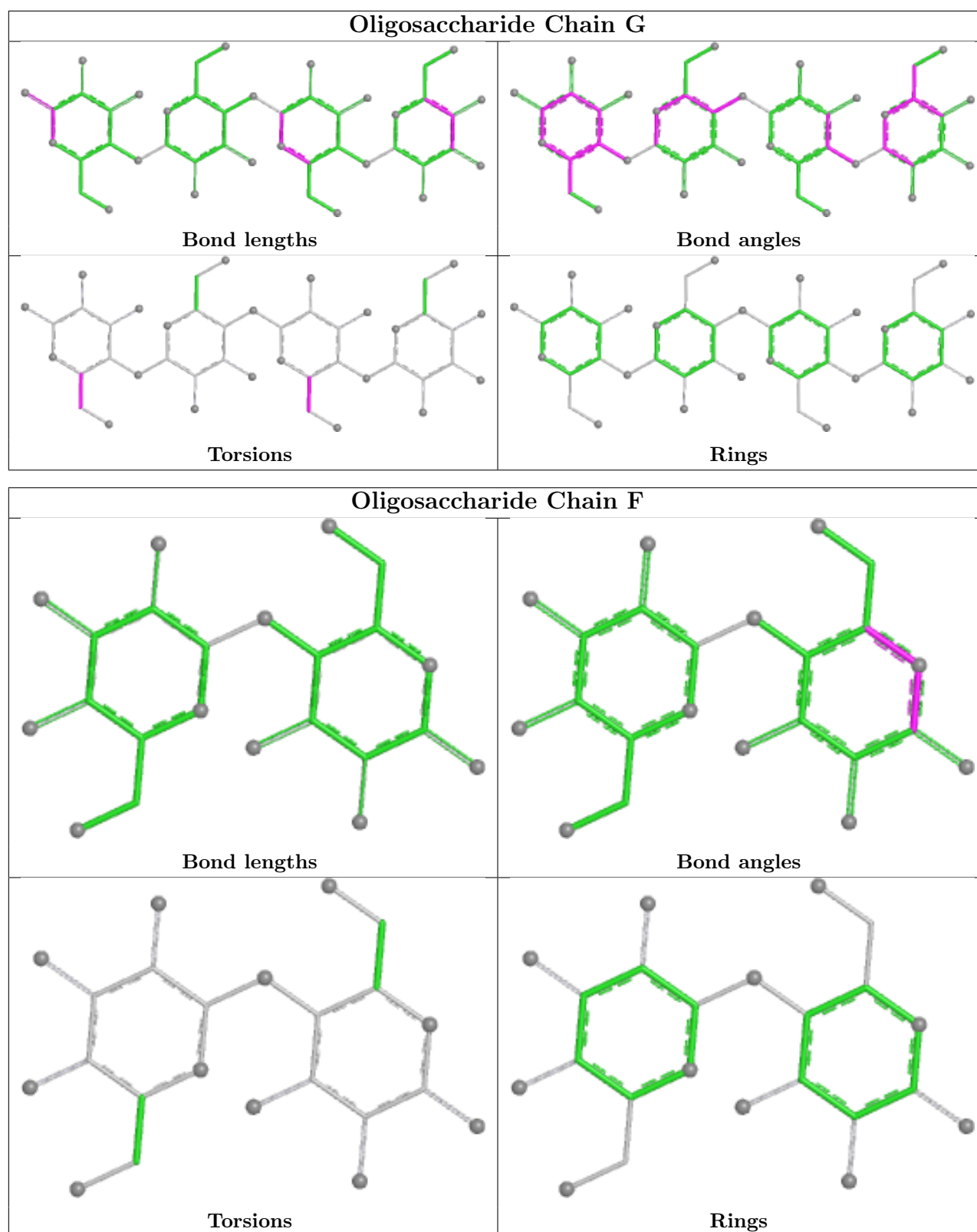
There are no ring outliers.

7 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	3	BGC	1	0
3	F	2	BGC	1	0
2	E	3	BGC	2	0
2	E	4	BGC	2	0
2	G	4	BGC	0	1
3	F	1	BGC	1	0
2	G	1	BGC	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](#)

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	377/458 (82%)	0.12	13 (3%) 48 47	6, 25, 49, 56	8 (2%)
1	B	377/458 (82%)	-0.17	10 (2%) 56 55	6, 22, 36, 46	10 (2%)
1	C	377/458 (82%)	0.78	31 (8%) 17 16	10, 38, 57, 64	9 (2%)
All	All	1131/1374 (82%)	0.24	54 (4%) 35 34	6, 27, 52, 64	27 (2%)

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	312	ALA	11.0
1	C	311	PRO	10.3
1	C	310	GLY	8.4
1	A	311	PRO	7.6
1	A	312	ALA	7.4

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.

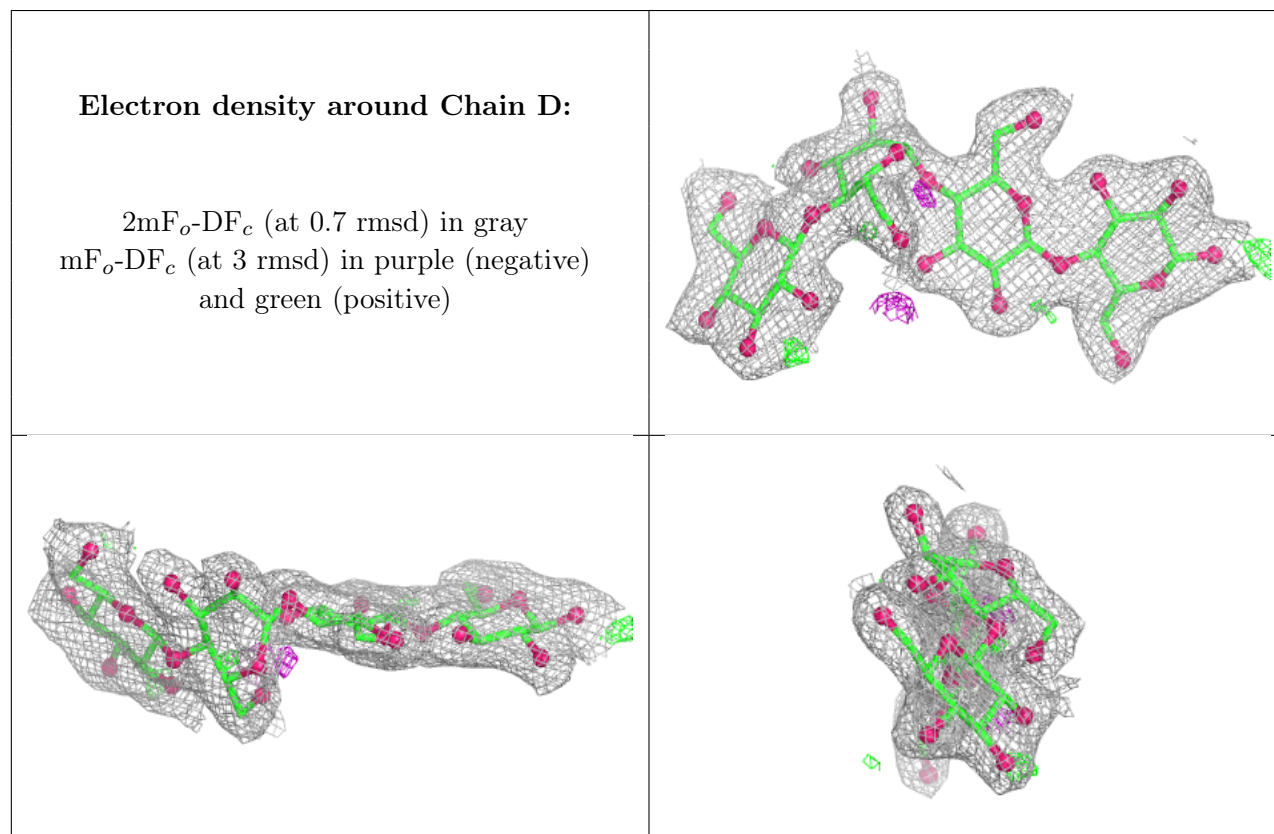
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	BGC	F	2	11/12	0.78	0.17	62,69,77,77	0
3	BGC	F	1	12/12	0.79	0.16	55,70,77,79	0
2	BGC	G	1	12/12	0.88	0.09	29,34,42,47	0

Continued on next page...

Continued from previous page...

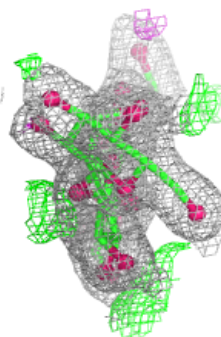
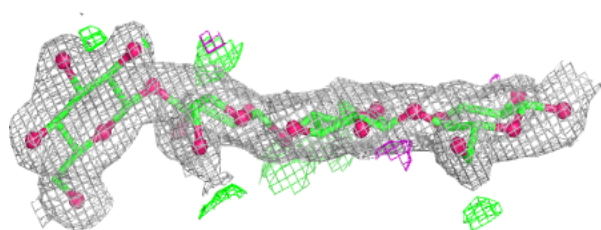
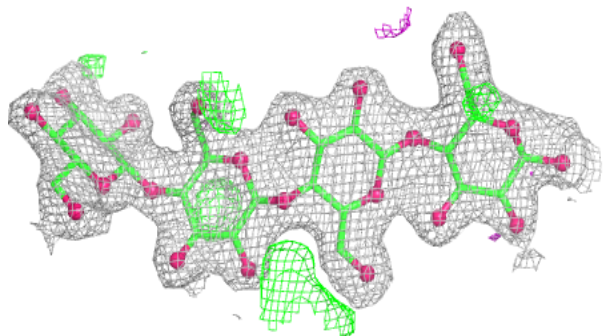
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	BGC	E	1	12/12	0.89	0.09	25,32,38,40	0
2	BGC	D	4	11/12	0.90	0.08	30,33,36,37	0
2	BGC	G	3	11/12	0.90	0.09	41,44,46,47	0
2	BGC	D	3	11/12	0.90	0.11	27,28,31,34	0
2	BGC	E	3	11/12	0.90	0.09	24,31,34,35	0
2	BGC	G	4	11/12	0.91	0.08	38,42,45,46	0
2	BGC	D	1	12/12	0.93	0.08	22,26,30,37	0
2	BGC	E	2	11/12	0.94	0.06	21,28,31,34	0
2	BGC	G	2	11/12	0.94	0.07	28,33,36,37	0
2	BGC	D	2	11/12	0.94	0.06	23,26,30,33	0
2	BGC	E	4	11/12	0.95	0.06	23,28,33,35	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.

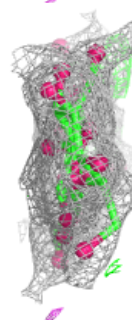
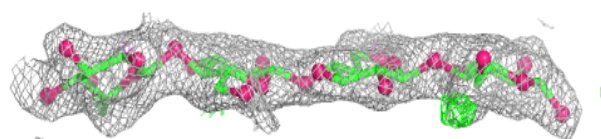
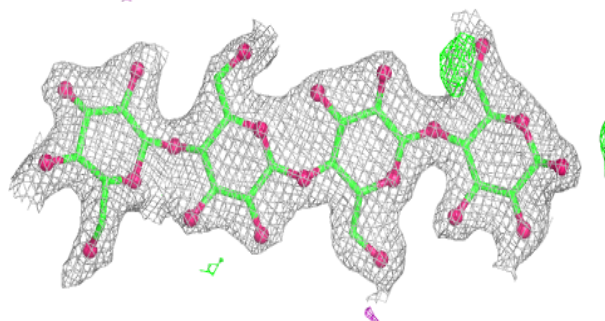


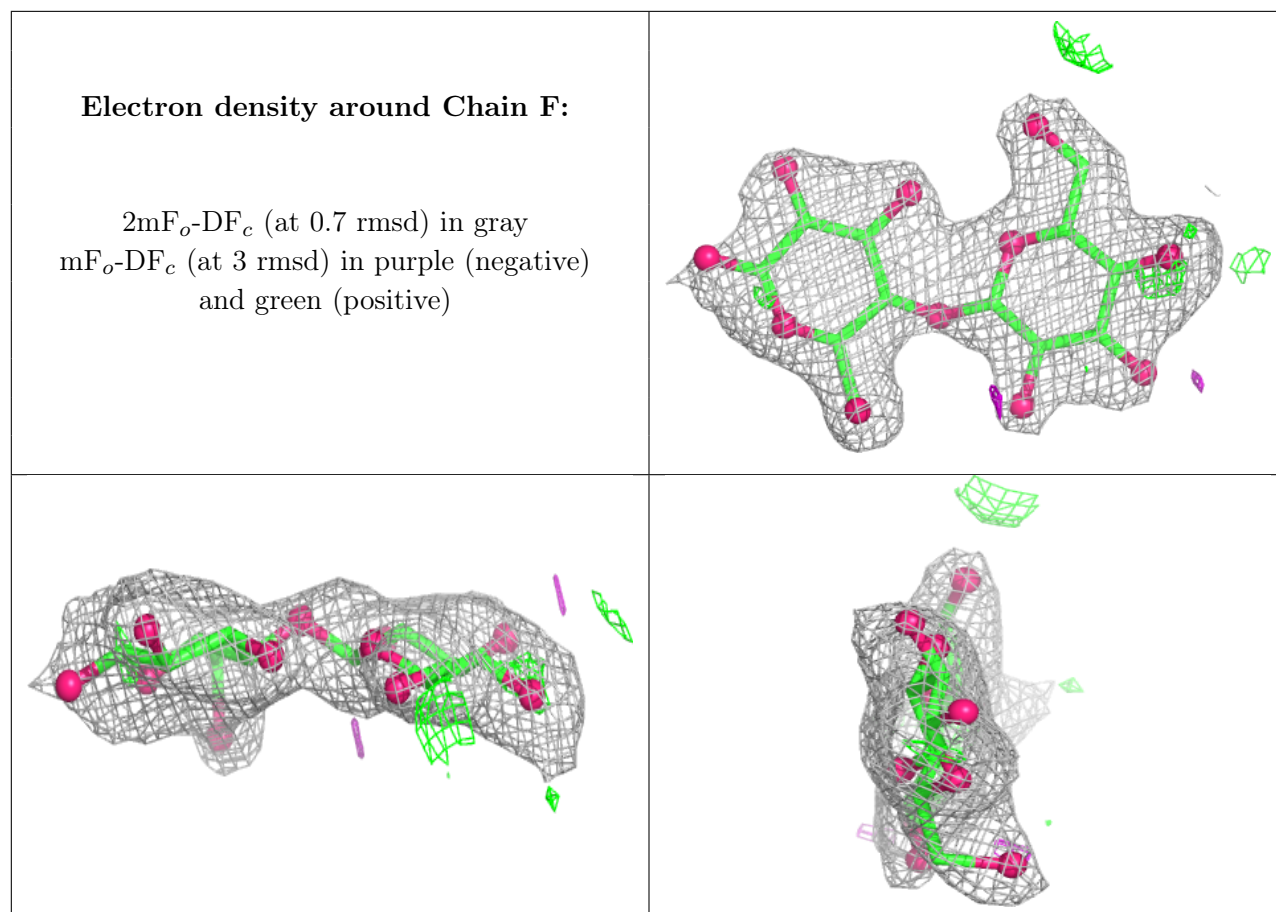
Electron density around Chain E:

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

**Electron density around Chain 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.