wwPDB X-ray Structure Validation Summary Report (i)

Feb 15, 2024 – 07:56 PM EST

PDB ID	:	3QXQ
Title	:	Structure of the bacterial cellulose synthase subunit Z in complex with cellopentaose
Authors	:	Zimmer, J.
Deposited on	:	2011-03-02
Resolution	:	2.20 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:

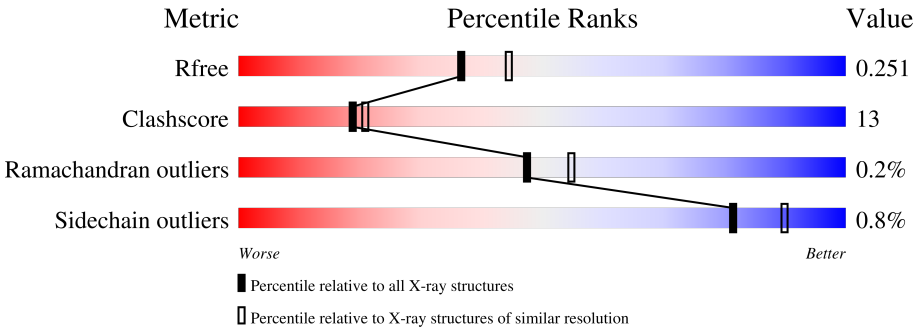
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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%

Mol	Chain	Length	Quality of chain
1	A	355	82% 13% 5%
1	B	355	79% 16% 5%
1	C	355	79% 17% 5%
1	D	355	70% 24% • 5%
2	E	5	20% 20% 60%
2	F	5	20% 40% 40%
2	G	5	60% 40%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	5	60% 40%

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	BGC	E	1	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 11761 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	338	Total	C	N	O	S	0	1	0
			2736	1750	476	500	10			
1	B	338	Total	C	N	O	S	0	0	0
			2728	1745	473	500	10			
1	C	338	Total	C	N	O	S	0	0	0
			2728	1745	473	500	10			
1	D	338	Total	C	N	O	S	0	0	0
			2728	1745	473	500	10			

There are 36 discrepancies between the modelled and reference sequences:

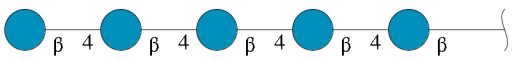
Chain	Residue	Modelled	Actual	Comment	Reference
A	55	GLN	GLU	engineered mutation	UNP P37651
A	369	LEU	-	expression tag	UNP P37651
A	370	GLU	-	expression tag	UNP P37651
A	371	HIS	-	expression tag	UNP P37651
A	372	HIS	-	expression tag	UNP P37651
A	373	HIS	-	expression tag	UNP P37651
A	374	HIS	-	expression tag	UNP P37651
A	375	HIS	-	expression tag	UNP P37651
A	376	HIS	-	expression tag	UNP P37651
B	55	GLN	GLU	engineered mutation	UNP P37651
B	369	LEU	-	expression tag	UNP P37651
B	370	GLU	-	expression tag	UNP P37651
B	371	HIS	-	expression tag	UNP P37651
B	372	HIS	-	expression tag	UNP P37651
B	373	HIS	-	expression tag	UNP P37651
B	374	HIS	-	expression tag	UNP P37651
B	375	HIS	-	expression tag	UNP P37651
B	376	HIS	-	expression tag	UNP P37651
C	55	GLN	GLU	engineered mutation	UNP P37651
C	369	LEU	-	expression tag	UNP P37651
C	370	GLU	-	expression tag	UNP P37651

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	371	HIS	-	expression tag	UNP P37651
C	372	HIS	-	expression tag	UNP P37651
C	373	HIS	-	expression tag	UNP P37651
C	374	HIS	-	expression tag	UNP P37651
C	375	HIS	-	expression tag	UNP P37651
C	376	HIS	-	expression tag	UNP P37651
D	55	GLN	GLU	engineered mutation	UNP P37651
D	369	LEU	-	expression tag	UNP P37651
D	370	GLU	-	expression tag	UNP P37651
D	371	HIS	-	expression tag	UNP P37651
D	372	HIS	-	expression tag	UNP P37651
D	373	HIS	-	expression tag	UNP P37651
D	374	HIS	-	expression tag	UNP P37651
D	375	HIS	-	expression tag	UNP P37651
D	376	HIS	-	expression tag	UNP P37651

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-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	E	5	Total C O 56 30 26	0	0	0
2	F	5	Total C O 56 30 26	0	0	0
2	G	5	Total C O 56 30 26	0	0	0
2	H	5	Total C O 56 30 26	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	195	Total O 195 195	0	0
3	B	185	Total O 185 185	0	0
3	C	162	Total O 162 162	0	0

Continued on next page...

Continued from previous page...

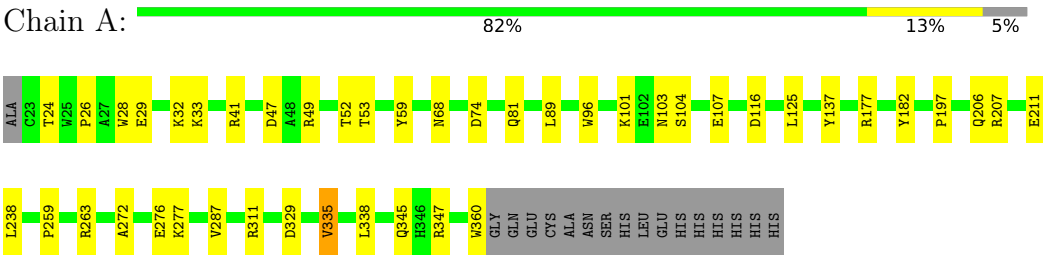
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	75	Total	O	0	0
			75	75		

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. 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. 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.

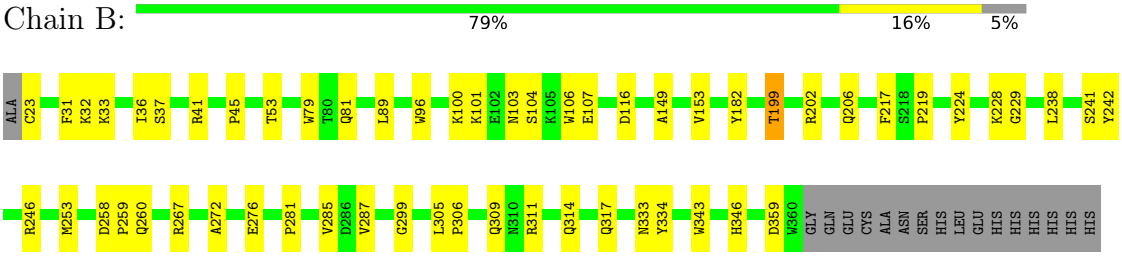
• Molecule 1: Endoglucanase

Chain A:



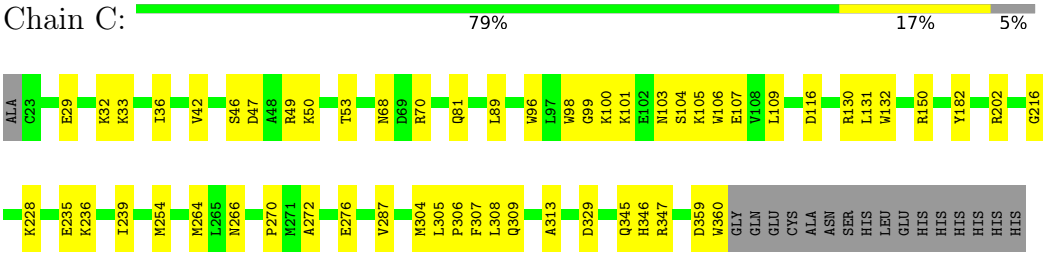
• Molecule 1: Endoglucanase

Chain B:



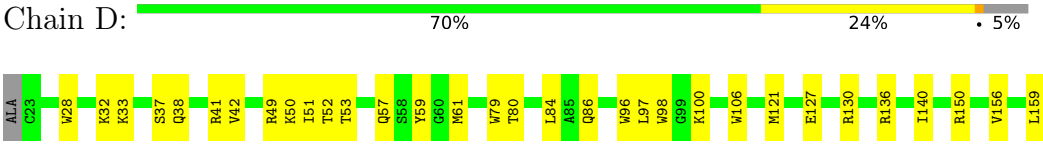
• Molecule 1: Endoglucanase

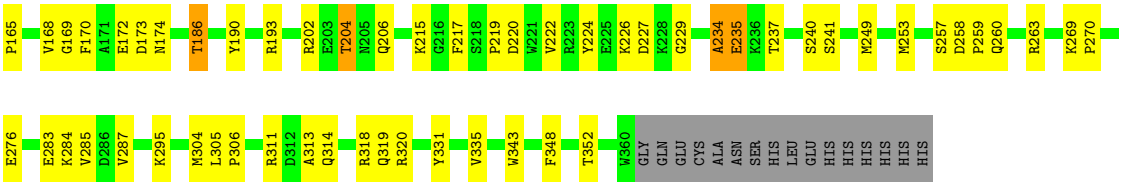
Chain C:



• Molecule 1: Endoglucanase

Chain D:



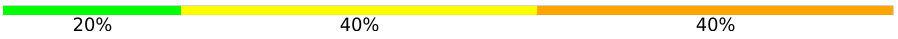


• Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain E:

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

Chain F:

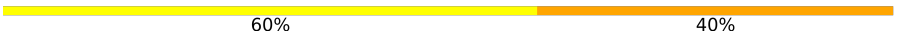


• Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain G:

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

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	90.68Å 99.61Å 93.07Å 90.00° 102.98° 90.00°	Depositor
Resolution (Å)	35.00 – 2.20 43.65 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.6 (35.00-2.20) 97.8 (43.65-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.82 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.168 , 0.227 0.205 , 0.251	Depositor DCC
R_{free} test set	5274 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	23.8	Xtriage
Anisotropy	0.908	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 34.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.018 for l,-k,h	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11761	wwPDB-VP
Average B, all atoms (Å^2)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.33% 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: 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	0.33	0/2819	0.46	0/3825
1	B	0.33	0/2808	0.47	0/3811
1	C	0.30	0/2808	0.44	0/3811
1	D	0.29	0/2808	0.43	0/3811
All	All	0.31	0/11243	0.45	0/15258

There are no bond length outliers.

There are no bond angle outliers.

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	2736	0	2657	56	0
1	B	2728	0	2644	49	0
1	C	2728	0	2644	66	0
1	D	2728	0	2644	113	0
2	E	56	0	48	9	0
2	F	56	0	48	5	0
2	G	56	0	48	10	0
2	H	56	0	48	9	0
3	A	195	0	0	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	185	0	0	5	0
3	C	162	0	0	5	0
3	D	75	0	0	8	0
All	All	11761	0	10781	287	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 13.

The worst 5 of 287 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:311[B]:ARG:HG3	1:B:311:ARG:NH1	1.52	1.24
1:C:103:ASN:HB2	1:C:105:LYS:H	1.01	1.13
1:A:101:LYS:HE3	1:A:107:GLU:HG3	1.17	1.09
2:H:5:BGC:H6C1	2:H:5:BGC:H2	1.37	1.05
1:B:96:TRP:CZ3	2:F:1:BGC:H6C1	1.93	1.04

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	337/355 (95%)	329 (98%)	8 (2%)	0	100 100
1	B	336/355 (95%)	330 (98%)	6 (2%)	0	100 100
1	C	336/355 (95%)	329 (98%)	7 (2%)	0	100 100
1	D	336/355 (95%)	327 (97%)	6 (2%)	3 (1%)	17 16
All	All	1345/1420 (95%)	1315 (98%)	27 (2%)	3 (0%)	47 55

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	234	ALA
1	D	86	GLN
1	D	235	GLU

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	284/297 (96%)	281 (99%)	3 (1%)	73	85
1	B	283/297 (95%)	281 (99%)	2 (1%)	84	91
1	C	283/297 (95%)	281 (99%)	2 (1%)	84	91
1	D	283/297 (95%)	281 (99%)	2 (1%)	84	91
All	All	1133/1188 (95%)	1124 (99%)	9 (1%)	81	90

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	186	THR
1	D	204	THR
1	B	23	CYS
1	B	199	THR
1	C	70	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	68	ASN
1	D	174	ASN
1	C	103	ASN
1	D	317	GLN
1	C	345	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)

20 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	E	1	2	12,12,12	0.49	0	17,17,17	1.33	2 (11%)
2	BGC	E	2	2	11,11,12	0.34	0	15,15,17	0.75	1 (6%)
2	BGC	E	3	2	11,11,12	0.23	0	15,15,17	0.64	0
2	BGC	E	4	2	11,11,12	0.21	0	15,15,17	0.65	0
2	BGC	E	5	2	11,11,12	0.28	0	15,15,17	1.26	2 (13%)
2	BGC	F	1	2	12,12,12	0.47	0	17,17,17	1.32	3 (17%)
2	BGC	F	2	2	11,11,12	0.28	0	15,15,17	0.74	1 (6%)
2	BGC	F	3	2	11,11,12	0.34	0	15,15,17	1.22	1 (6%)
2	BGC	F	4	2	11,11,12	0.21	0	15,15,17	0.70	0
2	BGC	F	5	2	11,11,12	0.29	0	15,15,17	1.03	1 (6%)
2	BGC	G	1	2	12,12,12	0.49	0	17,17,17	0.78	0
2	BGC	G	2	2	11,11,12	0.27	0	15,15,17	0.56	0
2	BGC	G	3	2	11,11,12	0.23	0	15,15,17	1.08	2 (13%)
2	BGC	G	4	2	11,11,12	0.25	0	15,15,17	1.39	2 (13%)
2	BGC	G	5	2	11,11,12	0.20	0	15,15,17	0.76	0
2	BGC	H	1	2	12,12,12	0.50	0	17,17,17	1.00	1 (5%)
2	BGC	H	2	2	11,11,12	0.23	0	15,15,17	1.23	2 (13%)
2	BGC	H	3	2	11,11,12	0.16	0	15,15,17	0.69	0
2	BGC	H	4	2	11,11,12	0.23	0	15,15,17	0.68	0
2	BGC	H	5	2	11,11,12	0.31	0	15,15,17	0.74	1 (6%)

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	E	1	2	-	0/2/22/22	0/1/1/1
2	BGC	E	2	2	-	0/2/19/22	0/1/1/1
2	BGC	E	3	2	-	0/2/19/22	0/1/1/1
2	BGC	E	4	2	-	0/2/19/22	0/1/1/1
2	BGC	E	5	2	-	0/2/19/22	0/1/1/1
2	BGC	F	1	2	-	0/2/22/22	0/1/1/1
2	BGC	F	2	2	-	0/2/19/22	0/1/1/1
2	BGC	F	3	2	-	2/2/19/22	0/1/1/1
2	BGC	F	4	2	-	2/2/19/22	0/1/1/1
2	BGC	F	5	2	-	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	-	2/2/19/22	0/1/1/1
2	BGC	G	4	2	-	2/2/19/22	0/1/1/1
2	BGC	G	5	2	-	2/2/19/22	0/1/1/1
2	BGC	H	1	2	-	2/2/22/22	0/1/1/1
2	BGC	H	2	2	-	0/2/19/22	0/1/1/1
2	BGC	H	3	2	-	2/2/19/22	0/1/1/1
2	BGC	H	4	2	-	0/2/19/22	0/1/1/1
2	BGC	H	5	2	-	1/2/19/22	0/1/1/1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	4	BGC	C1-O5-C5	3.74	117.25	112.19
2	E	5	BGC	C1-O5-C5	3.73	117.24	112.19
2	F	3	BGC	C1-O5-C5	3.69	117.19	112.19
2	E	1	BGC	C3-C4-C5	3.52	116.51	110.24
2	F	1	BGC	C3-C4-C5	3.35	116.22	110.24

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

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

Continued on next page...

Continued from previous page...

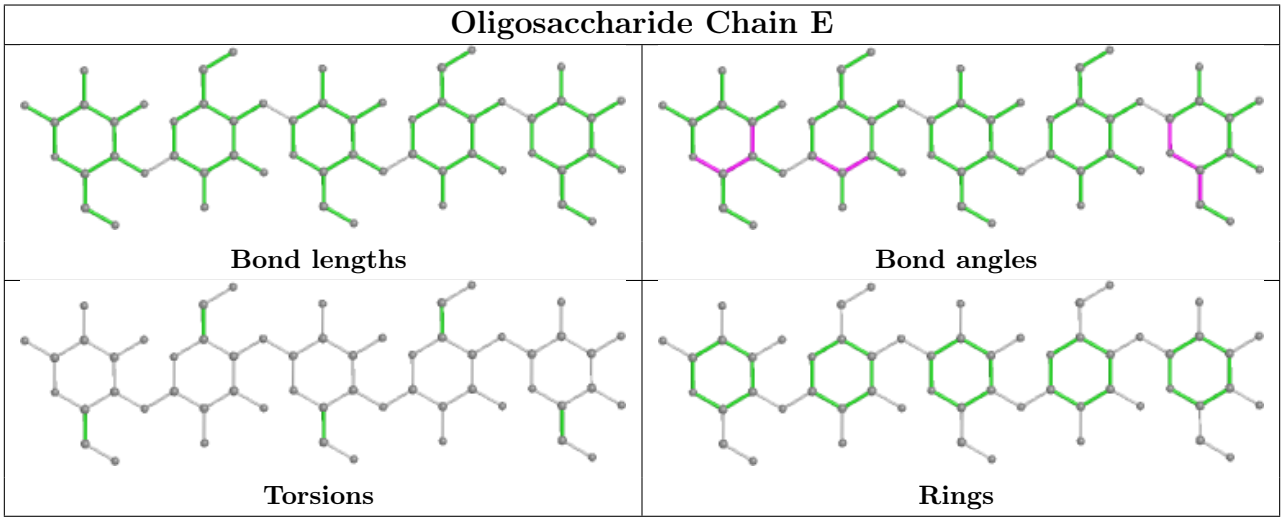
Mol	Chain	Res	Type	Atoms
2	F	3	BGC	O5-C5-C6-O6
2	G	4	BGC	O5-C5-C6-O6
2	G	4	BGC	C4-C5-C6-O6
2	F	3	BGC	C4-C5-C6-O6

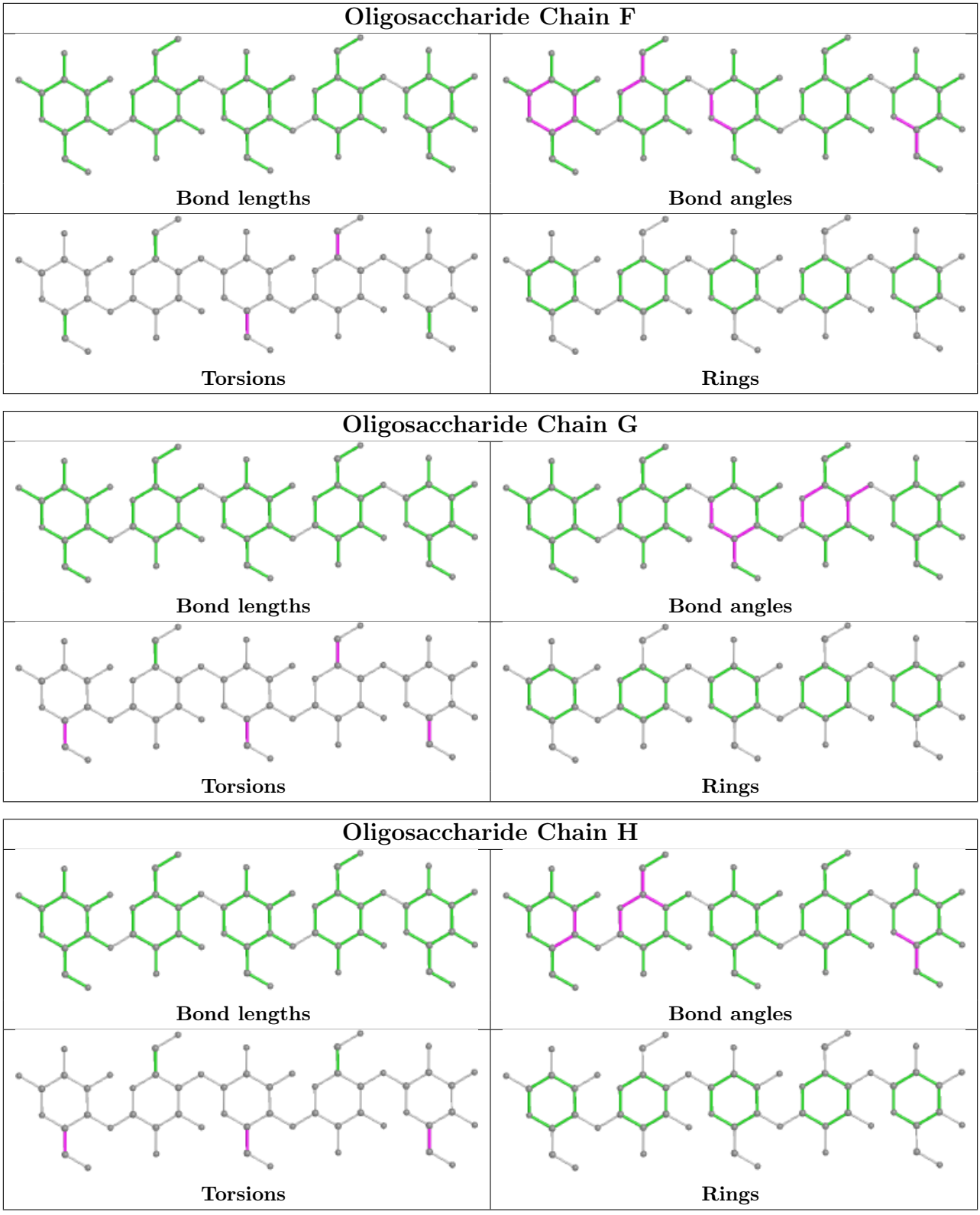
There are no ring outliers.

15 monomers are involved in 33 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	5	BGC	4	0
2	E	1	BGC	6	0
2	F	2	BGC	1	0
2	G	1	BGC	3	0
2	H	5	BGC	3	0
2	G	2	BGC	1	0
2	H	1	BGC	5	0
2	H	3	BGC	1	0
2	H	4	BGC	1	0
2	E	2	BGC	1	0
2	G	4	BGC	5	0
2	F	1	BGC	4	0
2	G	3	BGC	1	0
2	E	4	BGC	2	0
2	E	5	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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

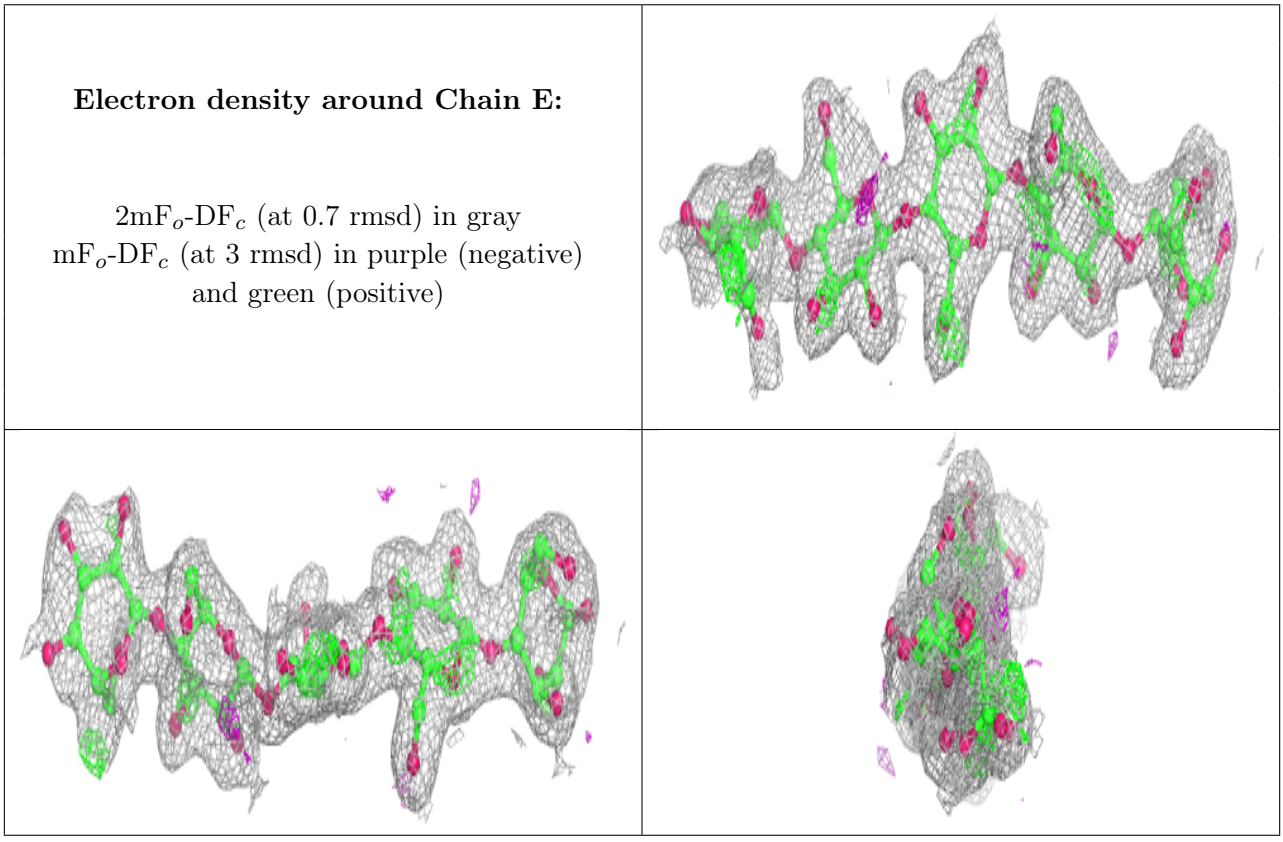
6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

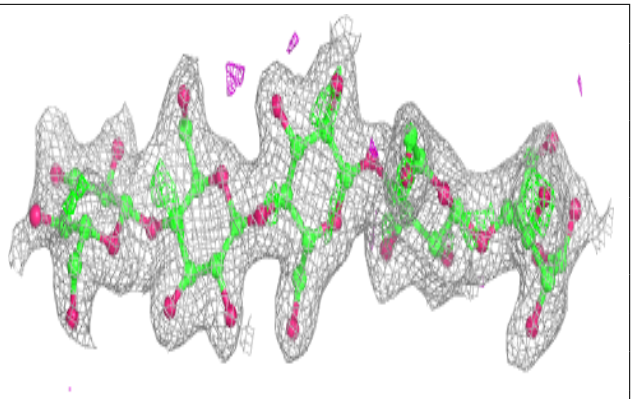
Unable to reproduce the depositors R factor - this section is therefore empty.

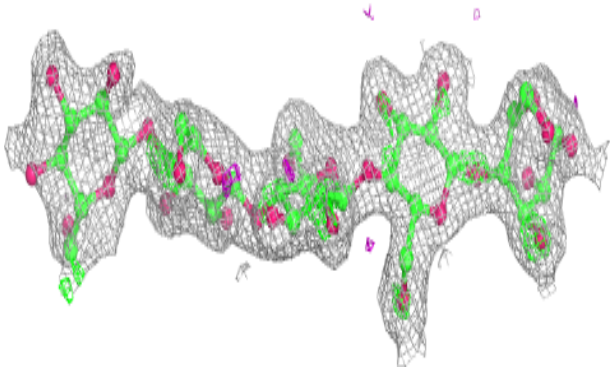
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

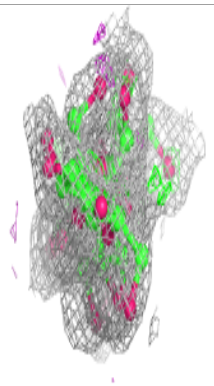


Electron density around Chain F:

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

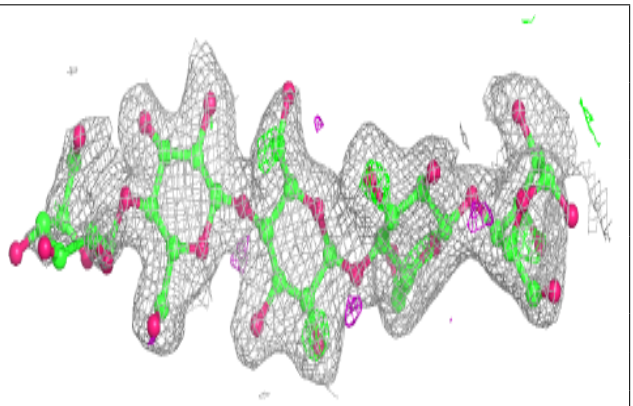


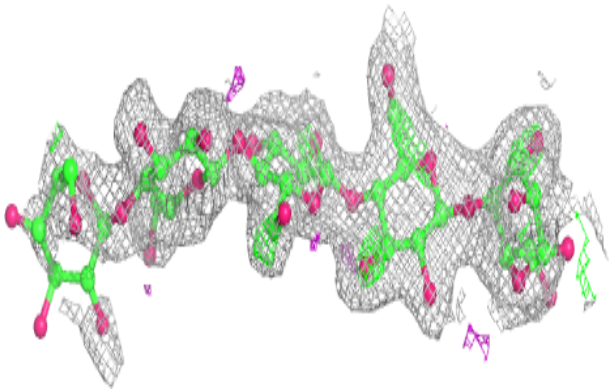


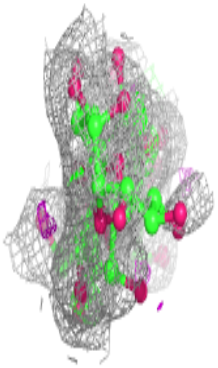


Electron density around Chain G:

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

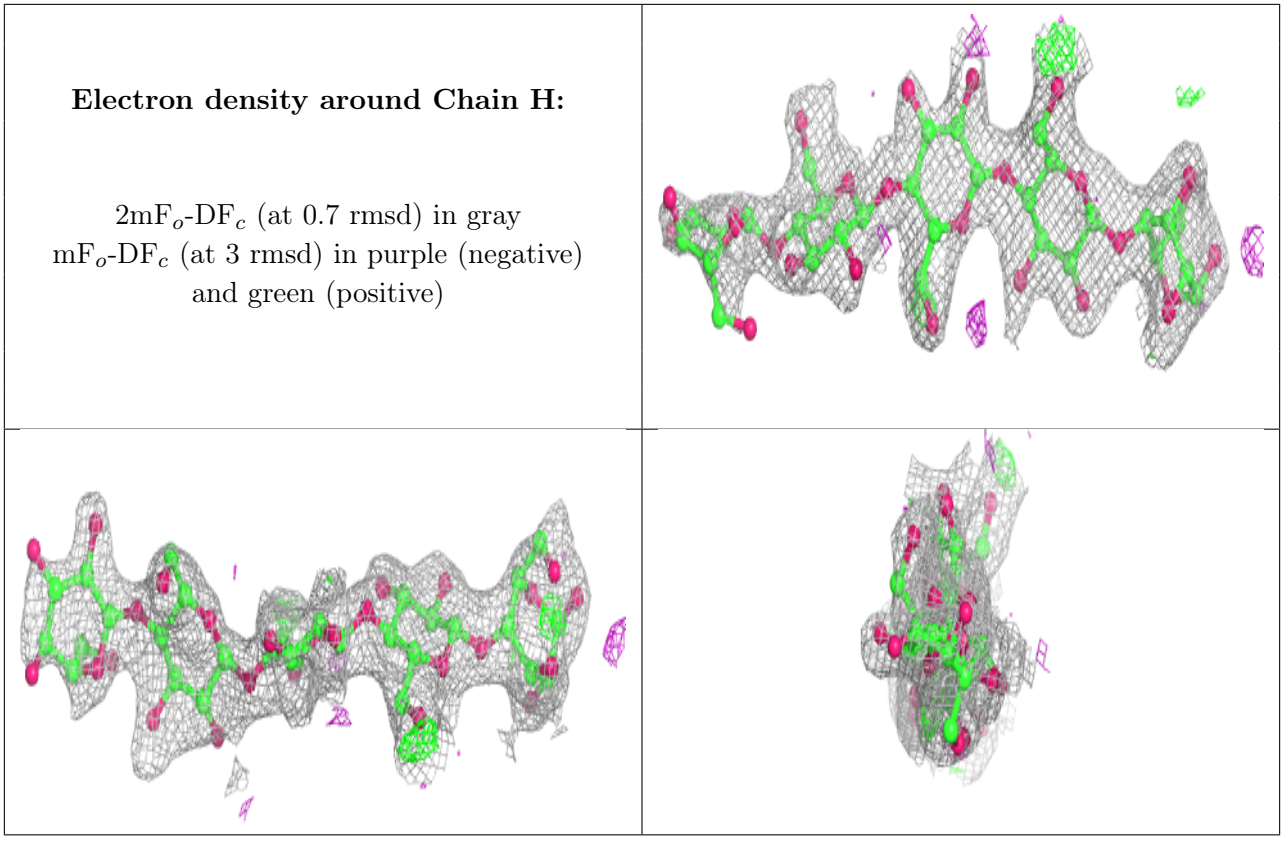






Electron density around Chain H:

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



6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.