



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2023 – 08:32 AM EDT

PDB ID : 7V0J
Title : Crystal structure of a CelR catalytic domain active site mutant with bound cellobiose product
Authors : Bingman, C.A.; Kuch, N.; Kutsche, M.E.; Parker, A.; Smith, R.W.; Fox, B.G.
Deposited on : 2022-05-10
Resolution : 2.40 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 : 4.02b-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

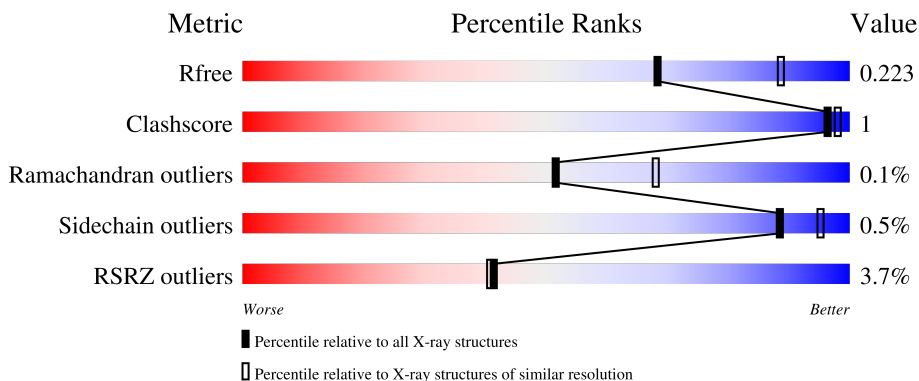
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	471	 3% 92%
1	B	471	 3% 94%
1	C	471	 5% 93%
2	D	2	 50% 50%
2	E	2	 100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	2	 100%

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 21272 atoms, of which 10189 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 Glucanase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	451	6949	2307	3355	595	674	18	0	0	0
1	B	453	7007	2329	3382	599	679	18	0	2	0
1	C	453	7008	2324	3386	601	679	18	0	2	0

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	18	MET	-	initiating methionine	UNP A3DCY5
A	439	GLN	GLU	conflict	UNP A3DCY5
A	473	ASN	-	expression tag	UNP A3DCY5
A	474	LEU	-	expression tag	UNP A3DCY5
A	475	TYR	-	expression tag	UNP A3DCY5
A	476	PHE	-	expression tag	UNP A3DCY5
A	477	GLN	-	expression tag	UNP A3DCY5
A	478	SER	-	expression tag	UNP A3DCY5
A	479	GLY	-	expression tag	UNP A3DCY5
A	480	SER	-	expression tag	UNP A3DCY5
A	481	HIS	-	expression tag	UNP A3DCY5
A	482	HIS	-	expression tag	UNP A3DCY5
A	483	HIS	-	expression tag	UNP A3DCY5
A	484	HIS	-	expression tag	UNP A3DCY5
A	485	HIS	-	expression tag	UNP A3DCY5
A	486	HIS	-	expression tag	UNP A3DCY5
A	487	HIS	-	expression tag	UNP A3DCY5
A	488	HIS	-	expression tag	UNP A3DCY5
B	18	MET	-	initiating methionine	UNP A3DCY5
B	439	GLN	GLU	conflict	UNP A3DCY5
B	473	ASN	-	expression tag	UNP A3DCY5
B	474	LEU	-	expression tag	UNP A3DCY5
B	475	TYR	-	expression tag	UNP A3DCY5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	476	PHE	-	expression tag	UNP A3DCY5
B	477	GLN	-	expression tag	UNP A3DCY5
B	478	SER	-	expression tag	UNP A3DCY5
B	479	GLY	-	expression tag	UNP A3DCY5
B	480	SER	-	expression tag	UNP A3DCY5
B	481	HIS	-	expression tag	UNP A3DCY5
B	482	HIS	-	expression tag	UNP A3DCY5
B	483	HIS	-	expression tag	UNP A3DCY5
B	484	HIS	-	expression tag	UNP A3DCY5
B	485	HIS	-	expression tag	UNP A3DCY5
B	486	HIS	-	expression tag	UNP A3DCY5
B	487	HIS	-	expression tag	UNP A3DCY5
B	488	HIS	-	expression tag	UNP A3DCY5
C	18	MET	-	initiating methionine	UNP A3DCY5
C	439	GLN	GLU	conflict	UNP A3DCY5
C	473	ASN	-	expression tag	UNP A3DCY5
C	474	LEU	-	expression tag	UNP A3DCY5
C	475	TYR	-	expression tag	UNP A3DCY5
C	476	PHE	-	expression tag	UNP A3DCY5
C	477	GLN	-	expression tag	UNP A3DCY5
C	478	SER	-	expression tag	UNP A3DCY5
C	479	GLY	-	expression tag	UNP A3DCY5
C	480	SER	-	expression tag	UNP A3DCY5
C	481	HIS	-	expression tag	UNP A3DCY5
C	482	HIS	-	expression tag	UNP A3DCY5
C	483	HIS	-	expression tag	UNP A3DCY5
C	484	HIS	-	expression tag	UNP A3DCY5
C	485	HIS	-	expression tag	UNP A3DCY5
C	486	HIS	-	expression tag	UNP A3DCY5
C	487	HIS	-	expression tag	UNP A3DCY5
C	488	HIS	-	expression tag	UNP A3DCY5

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	H	O			
2	D	2	45	12	22	11	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	2	Total	C	H	O	0	0	0
			45	12	22	11			
2	F	2	Total	C	H	O	0	0	0
			45	12	22	11			

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		
3	B	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		

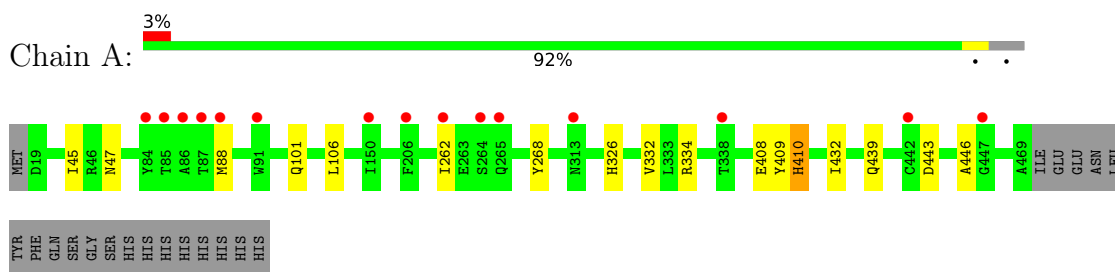
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	52	Total	O	0	0
			52	52		
4	B	50	Total	O	0	0
			50	50		
4	C	67	Total	O	0	1
			68	68		

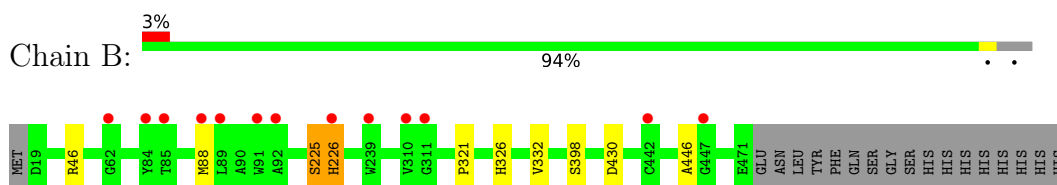
3 Residue-property plots [i](#)

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.

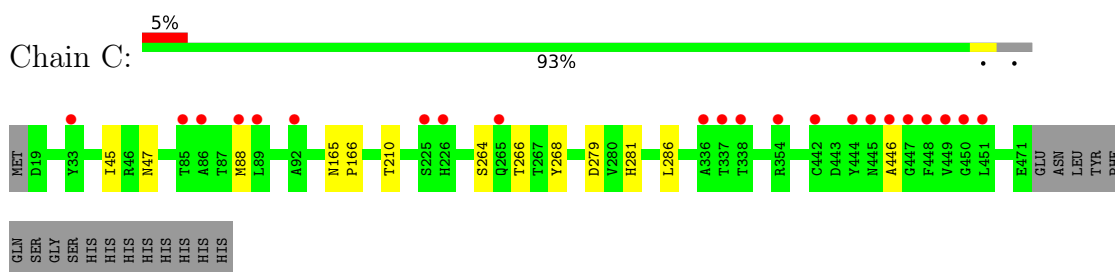
- Molecule 1: Glucanase



- Molecule 1: Glucanase



- Molecule 1: Glucanase



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



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



B6C1
B6C2

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

Chain F:

100%

B6C1
B6C2

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	92.58Å 161.64Å 210.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.17 – 2.40 40.17 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.8 (40.17-2.40) 99.8 (40.17-2.40)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.56 (at 2.39Å)	Xtrriage
Refinement program	PHENIX 1.20.1-4487	Depositor
R, R_{free}	0.191 , 0.224 0.190 , 0.223	Depositor DCC
R_{free} test set	3061 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å ²)	50.1	Xtrriage
Anisotropy	0.611	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 42.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.028 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.032 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	21272	wwPDB-VP
Average B, all atoms (Å ²)	70.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.01% 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: CA, 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	0.25	0/3711	0.45	0/5059
1	B	0.25	0/3750	0.45	0/5113
1	C	0.25	0/3750	0.45	0/5111
All	All	0.25	0/11211	0.45	0/15283

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 [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	3594	3355	3353	9	0
1	B	3625	3382	3375	5	0
1	C	3622	3386	3376	7	0
2	D	23	22	21	1	0
2	E	23	22	21	0	0
2	F	23	22	21	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	52	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	50	0	0	0	0
4	C	68	0	0	0	0
All	All	11083	10189	10167	20	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 1.

The worst 5 of 20 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:262:ILE:HG23	1:A:268:TYR:O	1.98	0.62
1:B:225:SER:O	1:B:226:HIS:HB2	2.05	0.56
1:A:45:ILE:HG22	1:A:47:ASN:H	1.72	0.54
1:C:279:ASP:OD1	1:C:281:HIS:ND1	2.40	0.53
1:C:45:ILE:HG22	1:C:47:ASN:H	1.76	0.50

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	449/471 (95%)	432 (96%)	17 (4%)	0	100	100
1	B	453/471 (96%)	433 (96%)	19 (4%)	1 (0%)	47	62
1	C	453/471 (96%)	436 (96%)	17 (4%)	0	100	100
All	All	1355/1413 (96%)	1301 (96%)	53 (4%)	1 (0%)	51	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	226	HIS

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	364/383 (95%)	363 (100%)	1 (0%)	92	97
1	B	368/383 (96%)	364 (99%)	4 (1%)	73	87
1	C	368/383 (96%)	367 (100%)	1 (0%)	92	97
All	All	1100/1149 (96%)	1094 (100%)	6 (0%)	88	95

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	398	SER
1	B	430	ASP
1	C	268	TYR
1	B	46	ARG
1	A	410	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

6 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	0.17	0	17,17,17	0.37	0
2	BGC	D	2	2	11,11,12	0.35	0	15,15,17	0.38	0
2	BGC	E	1	2	12,12,12	0.18	0	17,17,17	0.39	0
2	BGC	E	2	2	11,11,12	0.36	0	15,15,17	0.56	0
2	BGC	F	1	2	12,12,12	0.16	0	17,17,17	0.45	0
2	BGC	F	2	2	11,11,12	0.36	0	15,15,17	0.37	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	D	1	2	-	1/2/22/22	0/1/1/1
2	BGC	D	2	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	F	1	2	-	0/2/22/22	0/1/1/1
2	BGC	F	2	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

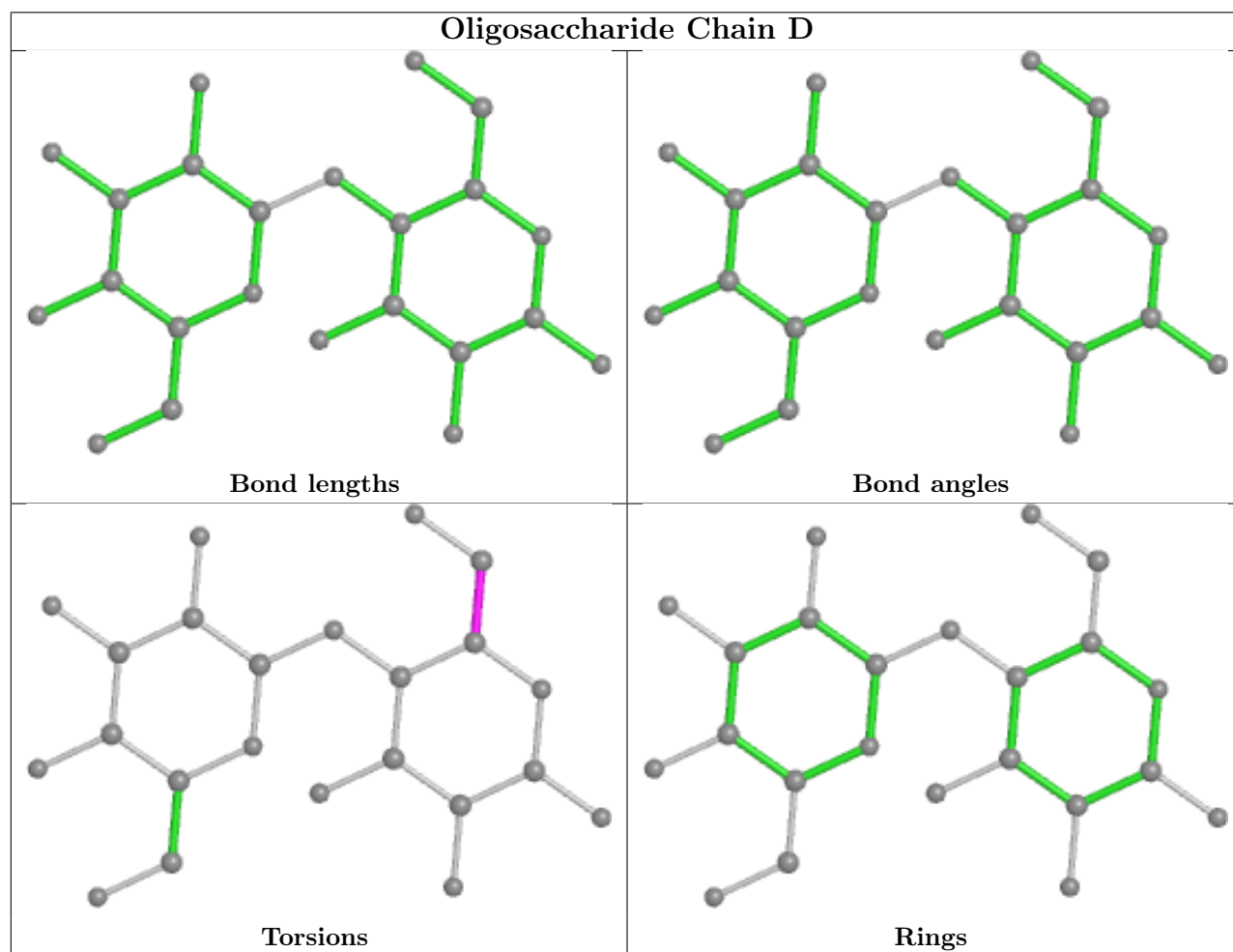
Mol	Chain	Res	Type	Atoms
2	E	1	BGC	C4-C5-C6-O6
2	E	1	BGC	O5-C5-C6-O6
2	D	1	BGC	O5-C5-C6-O6

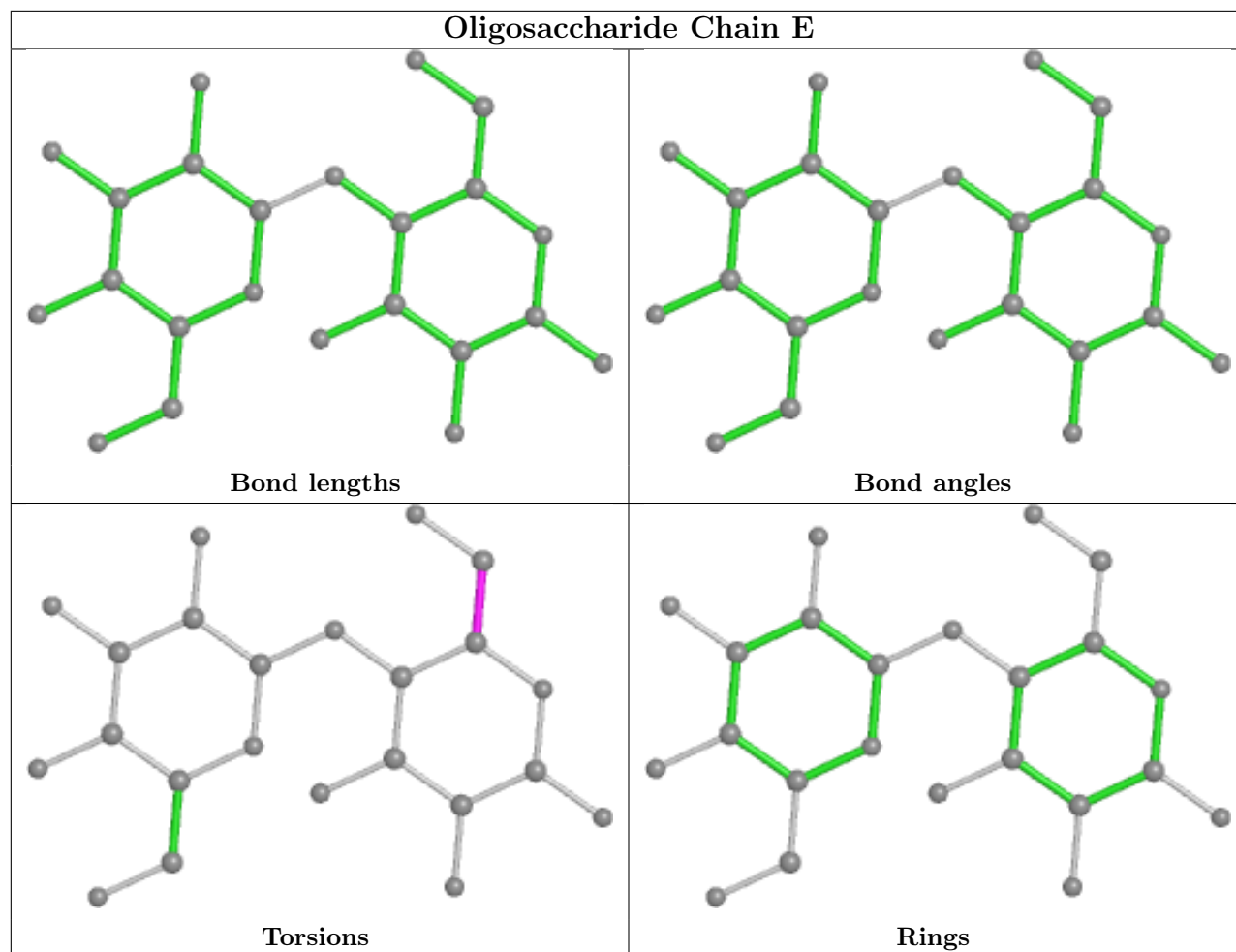
There are no ring outliers.

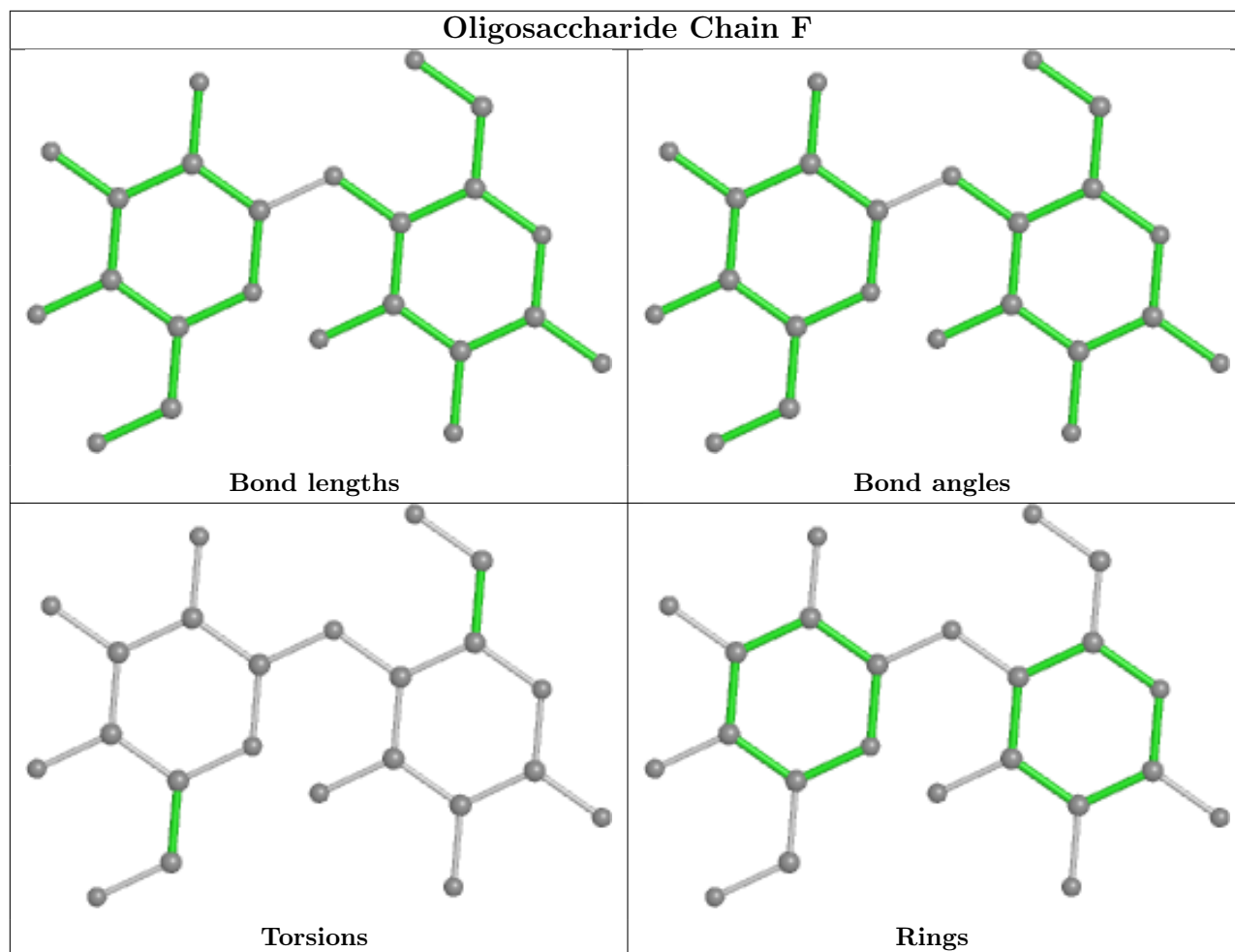
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	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](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

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 [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	451/471 (95%)	0.03	15 (3%) 46 45	48, 62, 82, 103	0
1	B	453/471 (96%)	0.00	13 (2%) 51 50	47, 65, 88, 115	0
1	C	453/471 (96%)	0.05	22 (4%) 29 28	46, 59, 80, 107	0
All	All	1357/1413 (96%)	0.03	50 (3%) 41 41	46, 62, 83, 115	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	447	GLY	4.1
1	C	85	THR	4.0
1	C	226	HIS	3.6
1	B	226	HIS	3.6
1	C	442	CYS	3.5

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.

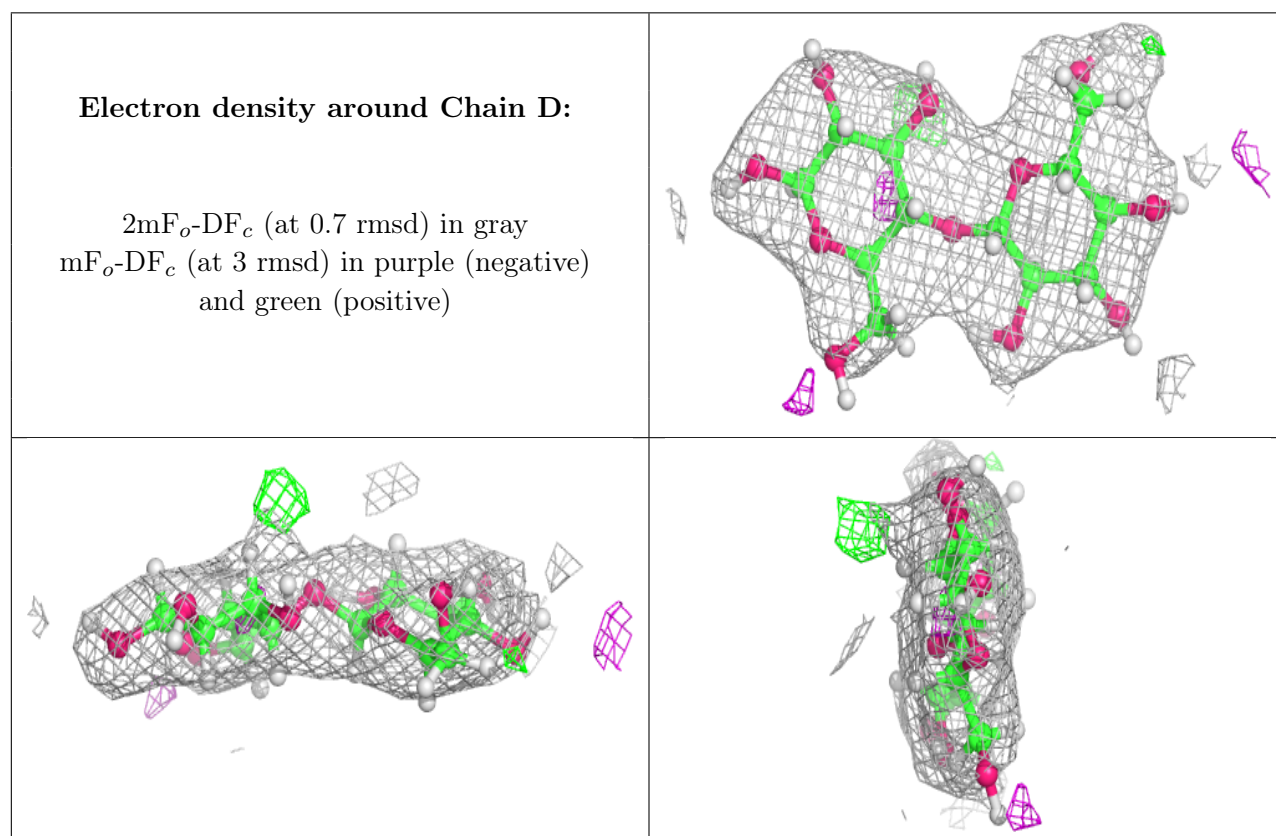
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	BGC	F	1	12/12	0.84	0.20	90,95,113,120	0
2	BGC	F	2	11/12	0.89	0.15	89,95,112,113	0
2	BGC	E	1	12/12	0.90	0.13	76,94,115,121	0

Continued on next page...

Continued from previous page...

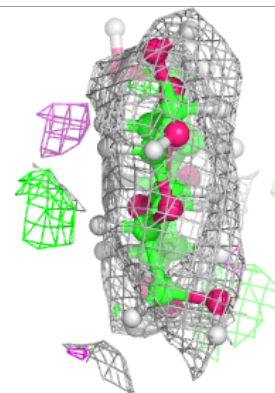
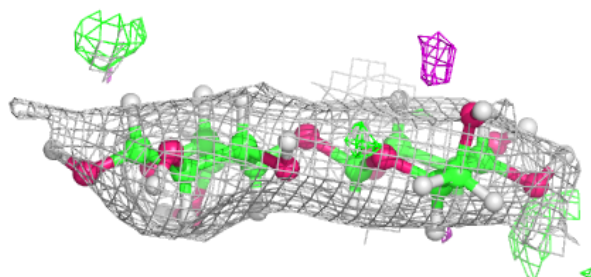
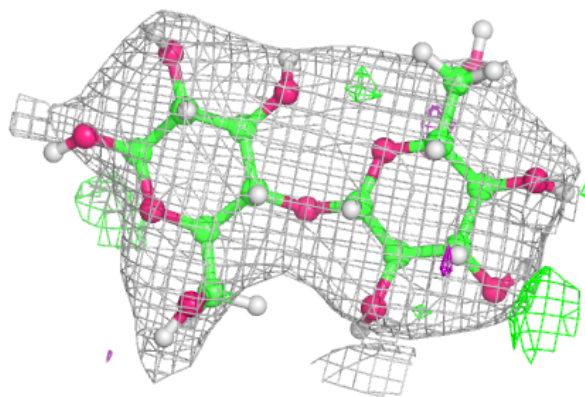
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	BGC	D	1	12/12	0.91	0.14	80,84,100,101	0
2	BGC	E	2	11/12	0.95	0.17	68,80,94,96	0
2	BGC	D	2	11/12	0.96	0.13	84,93,106,111	0

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.

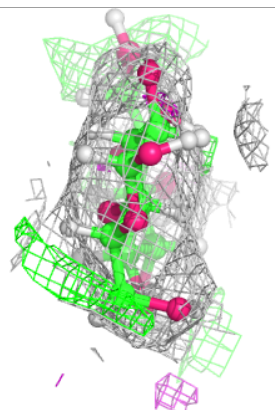
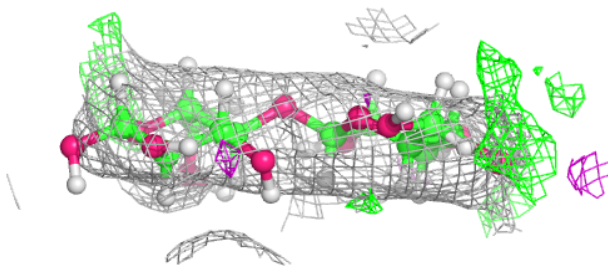
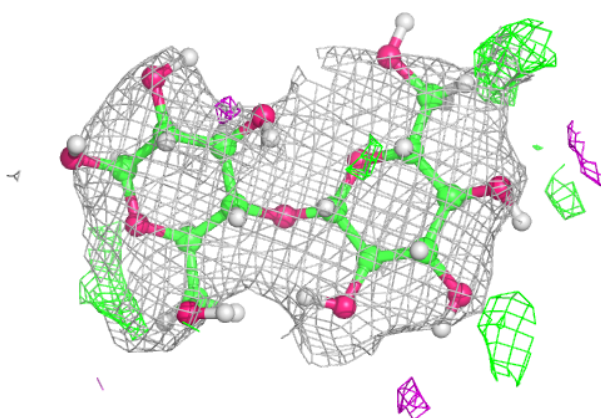


Electron density around Chain E:

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

**Electron density around Chain 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 [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	CA	B	500	1/1	0.92	0.07	75,75,75,75	0
3	CA	A	500	1/1	0.94	0.09	76,76,76,76	0
3	CA	C	500	1/1	0.94	0.10	71,71,71,71	0

6.5 Other polymers [i](#)

There are no such residues in this entry.