



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 28, 2023 – 03:40 AM EDT

PDB ID : 3KL5
Title : Structure Analysis of a Xylanase From Glycosyl Hydrolase Family Thirty: Carbohydrate Ligand Complexes Reveal this Family of Enzymes Unique Mechanism of Substrate Specificity and Recognition
Authors : St John, F.J.; Hurlbert, J.C.; Pozharski, E.
Deposited on : 2009-11-06
Resolution : 2.59 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 : 2.35
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) : 2.35

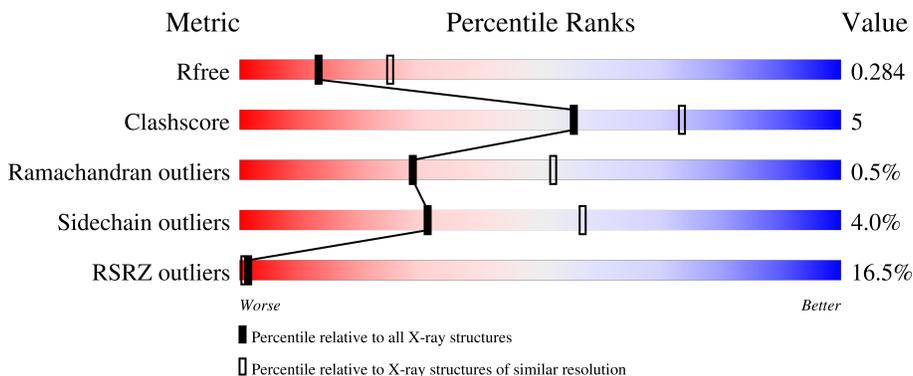
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.59 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (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	401	
1	B	401	
1	C	401	
1	D	401	
2	E	3	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	3	
2	G	3	
2	H	3	

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	XYP	H	1	-	-	-	X

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 12477 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 Glucuronoxylanase xynC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	389	3109	1966	552	582	9	0	1	0
1	B	390	3136	1984	557	586	9	0	3	0
1	C	389	3110	1967	552	582	9	0	1	0
1	D	367	2948	1878	516	545	9	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP Q45070
A	392	LEU	-	expression tag	UNP Q45070
A	393	GLU	-	expression tag	UNP Q45070
A	394	HIS	-	expression tag	UNP Q45070
A	395	HIS	-	expression tag	UNP Q45070
A	396	HIS	-	expression tag	UNP Q45070
A	397	HIS	-	expression tag	UNP Q45070
A	398	HIS	-	expression tag	UNP Q45070
A	399	HIS	-	expression tag	UNP Q45070
A	400	HIS	-	expression tag	UNP Q45070
A	401	HIS	-	expression tag	UNP Q45070
B	1	MET	-	expression tag	UNP Q45070
B	392	LEU	-	expression tag	UNP Q45070
B	393	GLU	-	expression tag	UNP Q45070
B	394	HIS	-	expression tag	UNP Q45070
B	395	HIS	-	expression tag	UNP Q45070
B	396	HIS	-	expression tag	UNP Q45070
B	397	HIS	-	expression tag	UNP Q45070
B	398	HIS	-	expression tag	UNP Q45070
B	399	HIS	-	expression tag	UNP Q45070
B	400	HIS	-	expression tag	UNP Q45070

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	401	HIS	-	expression tag	UNP Q45070
C	1	MET	-	expression tag	UNP Q45070
C	392	LEU	-	expression tag	UNP Q45070
C	393	GLU	-	expression tag	UNP Q45070
C	394	HIS	-	expression tag	UNP Q45070
C	395	HIS	-	expression tag	UNP Q45070
C	396	HIS	-	expression tag	UNP Q45070
C	397	HIS	-	expression tag	UNP Q45070
C	398	HIS	-	expression tag	UNP Q45070
C	399	HIS	-	expression tag	UNP Q45070
C	400	HIS	-	expression tag	UNP Q45070
C	401	HIS	-	expression tag	UNP Q45070
D	1	MET	-	expression tag	UNP Q45070
D	392	LEU	-	expression tag	UNP Q45070
D	393	GLU	-	expression tag	UNP Q45070
D	394	HIS	-	expression tag	UNP Q45070
D	395	HIS	-	expression tag	UNP Q45070
D	396	HIS	-	expression tag	UNP Q45070
D	397	HIS	-	expression tag	UNP Q45070
D	398	HIS	-	expression tag	UNP Q45070
D	399	HIS	-	expression tag	UNP Q45070
D	400	HIS	-	expression tag	UNP Q45070
D	401	HIS	-	expression tag	UNP Q45070

- Molecule 2 is an oligosaccharide called 4-O-methyl-alpha-D-glucofuranuronic acid-(1-2)-beta-D-xylofuranose-(1-4)-beta-D-xylofuranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	E	3	Total	C	O	0	0	0
			32	17	15			
2	F	3	Total	C	O	0	0	0
			32	17	15			
2	G	3	Total	C	O	0	0	0
			31	17	14			
2	H	3	Total	C	O	0	0	0
			32	17	15			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	10	Total 10	O 10	0	0
3	B	17	Total 17	O 17	0	0
3	C	19	Total 19	O 19	0	0
3	D	1	Total 1	O 1	0	0

XYP1
XYP2
GCV3

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	137.72Å 194.01Å 65.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.59 44.67 – 2.59	Depositor EDS
% Data completeness (in resolution range)	97.5 (50.00-2.59) 97.6 (44.67-2.59)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.32 (at 2.58Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.241 , 0.290 0.237 , 0.284	Depositor DCC
R_{free} test set	2778 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	65.9	Xtrriage
Anisotropy	0.068	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 43.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	12477	wwPDB-VP
Average B, all atoms (Å ²)	73.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.68% 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: GCV, XYP

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	0.59	0/3196	0.63	4/4353 (0.1%)
1	B	0.58	0/3224	0.70	5/4390 (0.1%)
1	C	0.59	2/3197 (0.1%)	0.70	5/4354 (0.1%)
1	D	0.43	0/3031	0.51	0/4123
All	All	0.55	2/12648 (0.0%)	0.64	14/17220 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	162	ASN	CG-OD1	-5.33	1.12	1.24
1	C	162	ASN	CG-ND2	-5.22	1.19	1.32

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	169	ARG	NE-CZ-NH1	-14.19	113.21	120.30
1	C	60	ARG	NE-CZ-NH2	12.89	126.75	120.30
1	C	60	ARG	NE-CZ-NH1	-12.73	113.93	120.30
1	B	169	ARG	NE-CZ-NH2	12.08	126.34	120.30
1	C	169	ARG	NE-CZ-NH2	-8.86	115.87	120.30

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	3109	0	2963	24	0
1	B	3136	0	2993	26	0
1	C	3110	0	2965	33	0
1	D	2948	0	2810	56	0
2	E	32	0	9	0	0
2	F	32	0	9	2	0
2	G	31	0	9	0	0
2	H	32	0	9	0	0
3	A	10	0	0	2	0
3	B	17	0	0	0	0
3	C	19	0	0	4	0
3	D	1	0	0	0	0
All	All	12477	0	11767	124	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 5.

The worst 5 of 124 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:C:254:HIS:HD2	3:C:411:HOH:O	1.57	0.87
1:B:244[B]:PRO:O	1:B:245[B]:GLU:HB2	1.76	0.82
1:B:244[B]:PRO:O	1:B:245[B]:GLU:CB	2.30	0.76
1:D:144:ALA:HB3	1:D:149:TRP:HD1	1.53	0.74
1:B:315:TYR:OH	1:C:222:GLY:HA2	1.86	0.74

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	388/401 (97%)	373 (96%)	15 (4%)	0	100	100
1	B	391/401 (98%)	373 (95%)	15 (4%)	3 (1%)	19	39
1	C	388/401 (97%)	373 (96%)	14 (4%)	1 (0%)	41	64
1	D	357/401 (89%)	316 (88%)	36 (10%)	5 (1%)	11	22
All	All	1524/1604 (95%)	1435 (94%)	80 (5%)	9 (1%)	29	47

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	268	TRP
1	B	245[A]	GLU
1	B	245[B]	GLU
1	B	4	ASP
1	D	58	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	331/341 (97%)	321 (97%)	10 (3%)	41	67
1	B	334/341 (98%)	321 (96%)	13 (4%)	32	58
1	C	331/341 (97%)	320 (97%)	11 (3%)	38	64
1	D	312/341 (92%)	293 (94%)	19 (6%)	18	38
All	All	1308/1364 (96%)	1255 (96%)	53 (4%)	31	56

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	213	TYR
1	D	60	ARG
1	D	307	THR
1	C	230	VAL
1	C	337	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	153	GLN
1	D	309	ASN
1	D	363	GLN
1	D	313	ASN
1	D	24	HIS

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](#)

12 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	XYP	E	1	2	10,10,10	1.02	1 (10%)	14,14,14	1.05	1 (7%)
2	XYP	E	2	2	9,9,10	1.81	4 (44%)	10,12,14	2.08	3 (30%)
2	GCV	E	3	2	13,13,14	1.53	2 (15%)	14,18,20	1.11	0
2	XYP	F	1	2	10,10,10	1.20	2 (20%)	14,14,14	1.63	2 (14%)
2	XYP	F	2	2	9,9,10	0.88	0	10,12,14	1.67	1 (10%)
2	GCV	F	3	2	13,13,14	0.85	0	14,18,20	0.83	0
2	XYP	G	1	2	9,9,10	1.46	1 (11%)	10,12,14	1.91	2 (20%)
2	XYP	G	2	2	9,9,10	1.95	3 (33%)	10,12,14	2.10	2 (20%)
2	GCV	G	3	2	13,13,14	1.73	3 (23%)	14,18,20	1.36	3 (21%)
2	XYP	H	1	2	10,10,10	2.18	5 (50%)	14,14,14	1.89	2 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	XYP	H	2	2	9,9,10	1.93	3 (33%)	10,12,14	3.04	2 (20%)
2	GCV	H	3	2	13,13,14	1.62	2 (15%)	14,18,20	2.15	3 (21%)

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	XYP	E	1	2	-	-	0/1/1/1
2	XYP	E	2	2	-	-	0/1/1/1
2	GCV	E	3	2	-	1/6/23/26	0/1/1/1
2	XYP	F	1	2	-	-	0/1/1/1
2	XYP	F	2	2	-	-	0/1/1/1
2	GCV	F	3	2	-	0/6/23/26	0/1/1/1
2	XYP	G	1	2	-	-	0/1/1/1
2	XYP	G	2	2	-	-	0/1/1/1
2	GCV	G	3	2	-	0/6/23/26	0/1/1/1
2	XYP	H	1	2	-	-	0/1/1/1
2	XYP	H	2	2	-	-	0/1/1/1
2	GCV	H	3	2	-	0/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	2	XYP	O5-C1	3.86	1.50	1.42
2	G	3	GCV	C4-C5	3.76	1.59	1.52
2	H	3	GCV	O5-C1	3.53	1.49	1.43
2	H	1	XYP	C4-C3	3.22	1.57	1.52
2	H	1	XYP	O4-C4	3.20	1.50	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	2	XYP	C1-C2-C3	-7.03	101.02	109.67
2	H	3	GCV	O5-C1-C2	5.83	119.77	110.77
2	H	2	XYP	O2-C2-C1	5.52	120.45	109.15
2	H	1	XYP	O5-C5-C4	5.03	118.54	110.77
2	G	1	XYP	C5-C4-C3	-5.01	103.50	109.67

There are no chirality outliers.

All (1) torsion outliers are listed below:

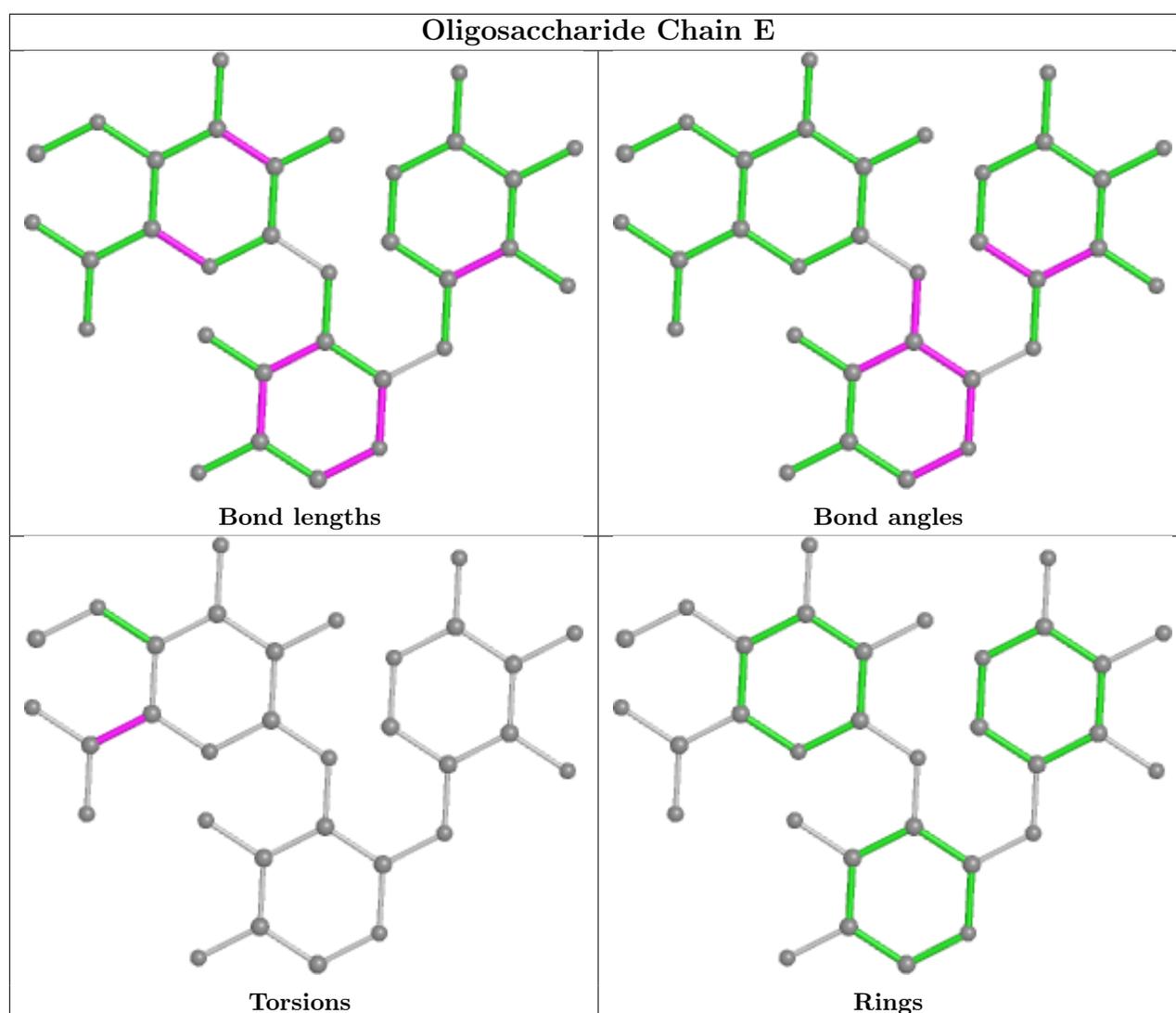
Mol	Chain	Res	Type	Atoms
2	E	3	GCV	O5-C5-C6-O6B

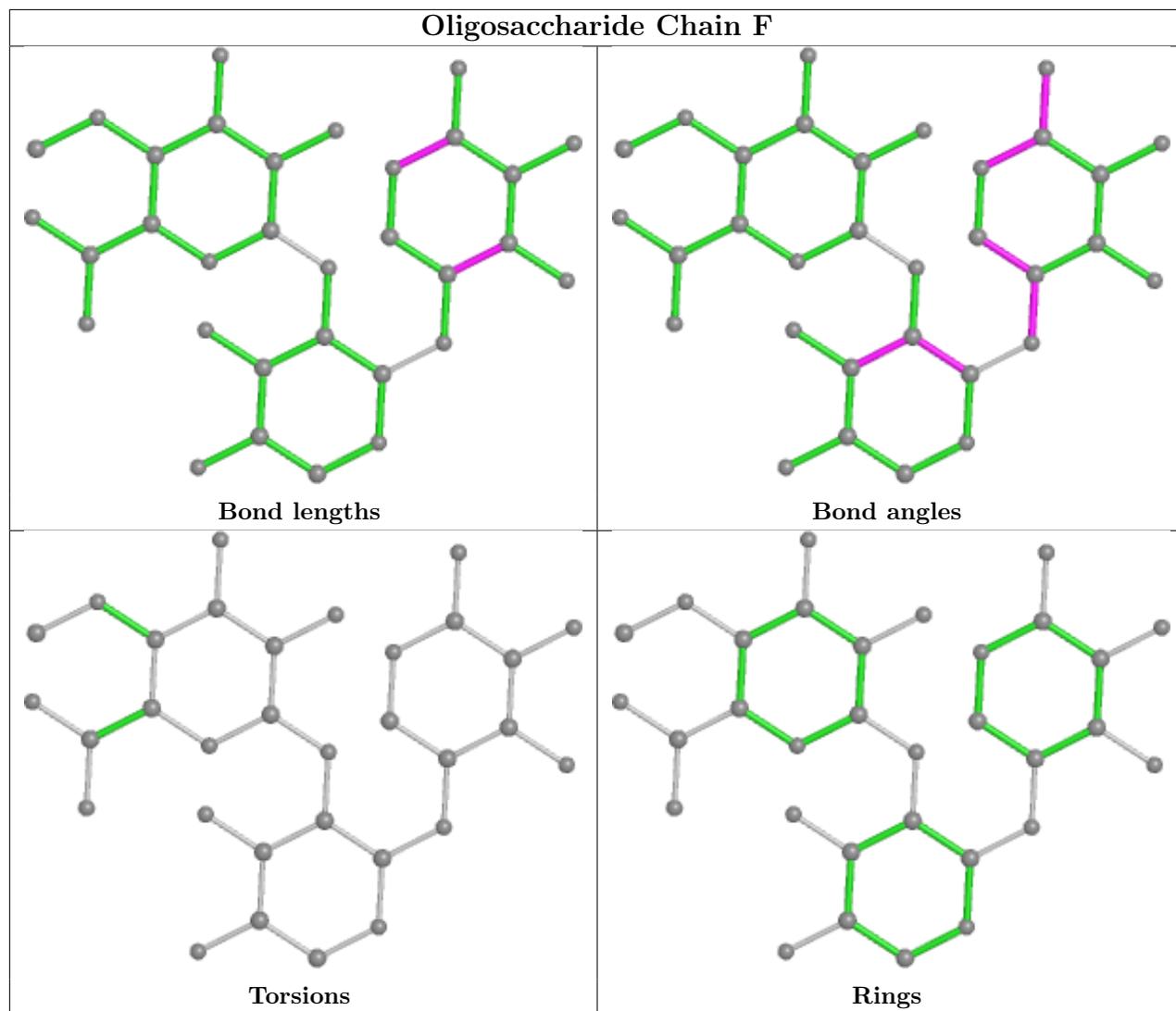
There are no ring outliers.

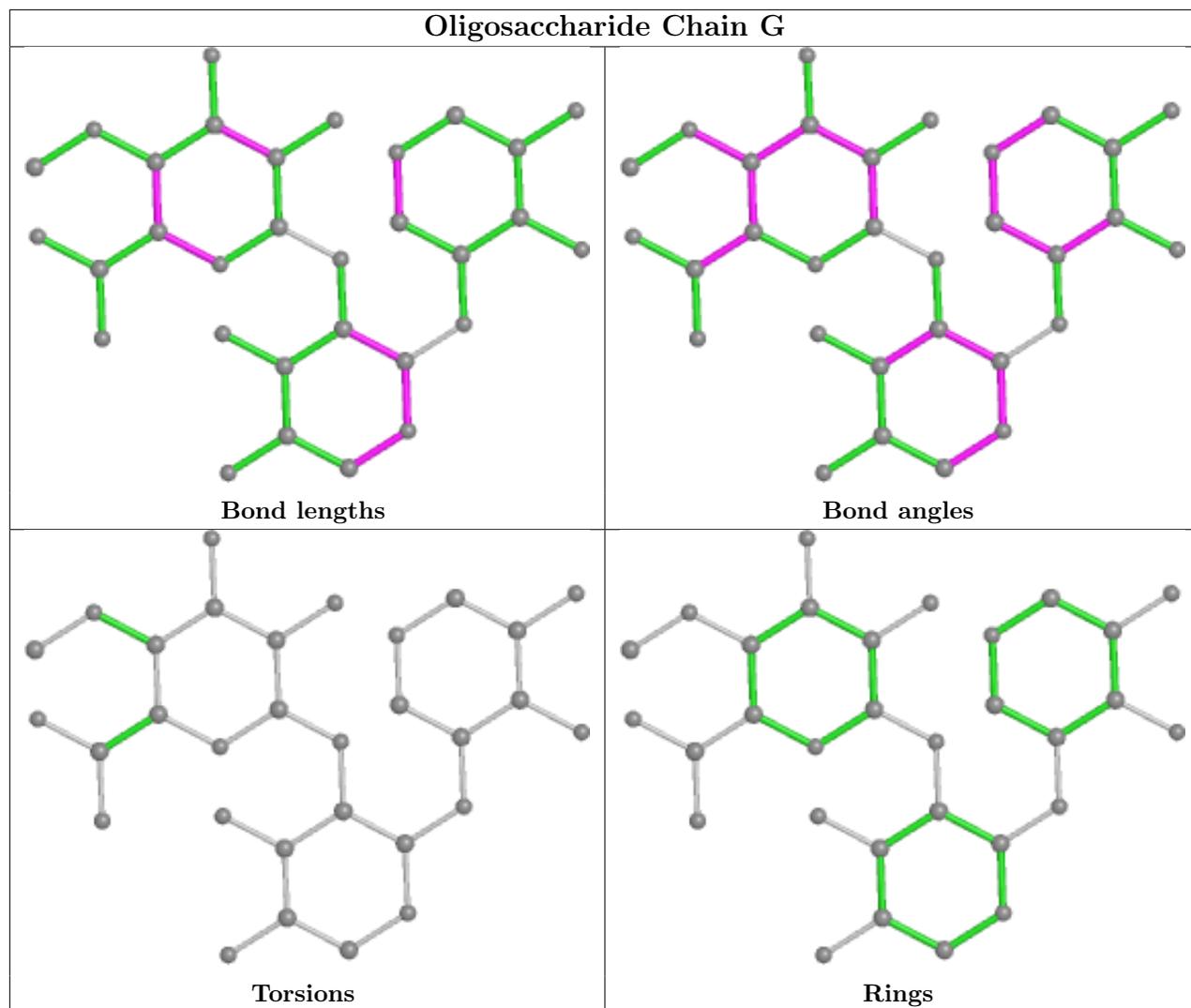
1 monomer is involved in 2 short contacts:

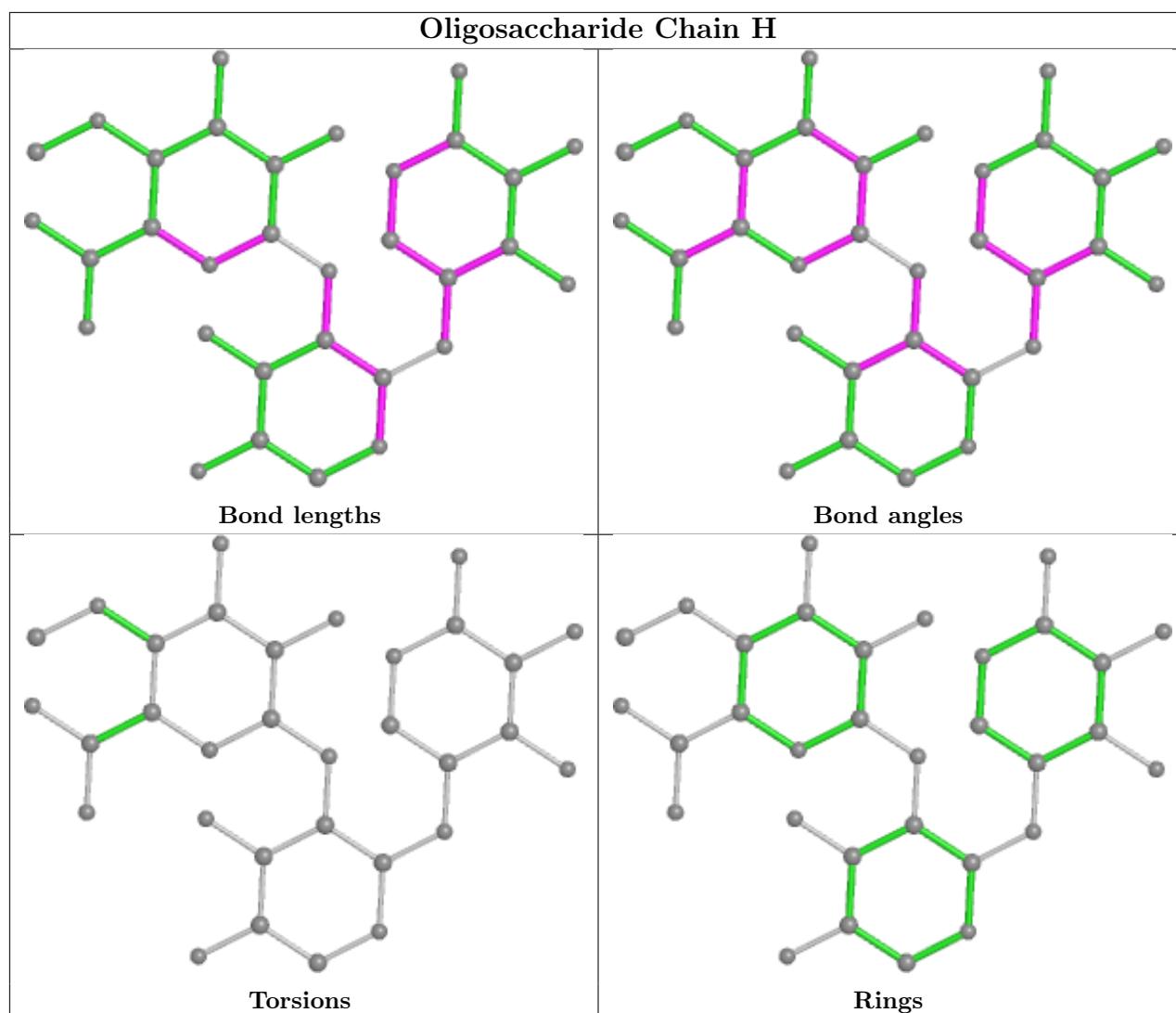
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	1	XYP	2	0

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	389/401 (97%)	0.41	25 (6%) 19 14	42, 58, 72, 80	0
1	B	390/401 (97%)	0.42	6 (1%) 73 70	42, 58, 73, 79	1 (0%)
1	C	389/401 (97%)	0.49	8 (2%) 63 58	42, 58, 72, 79	0
1	D	367/401 (91%)	2.75	215 (58%) 0 0	74, 120, 155, 179	364 (99%)
All	All	1535/1604 (95%)	0.99	254 (16%) 1 1	42, 62, 140, 179	365 (23%)

The worst 5 of 254 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	102	SER	10.0
1	D	343	LEU	9.6
1	D	100	ASP	9.2
1	D	103	ALA	8.9
1	D	319	TYR	8.0

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.

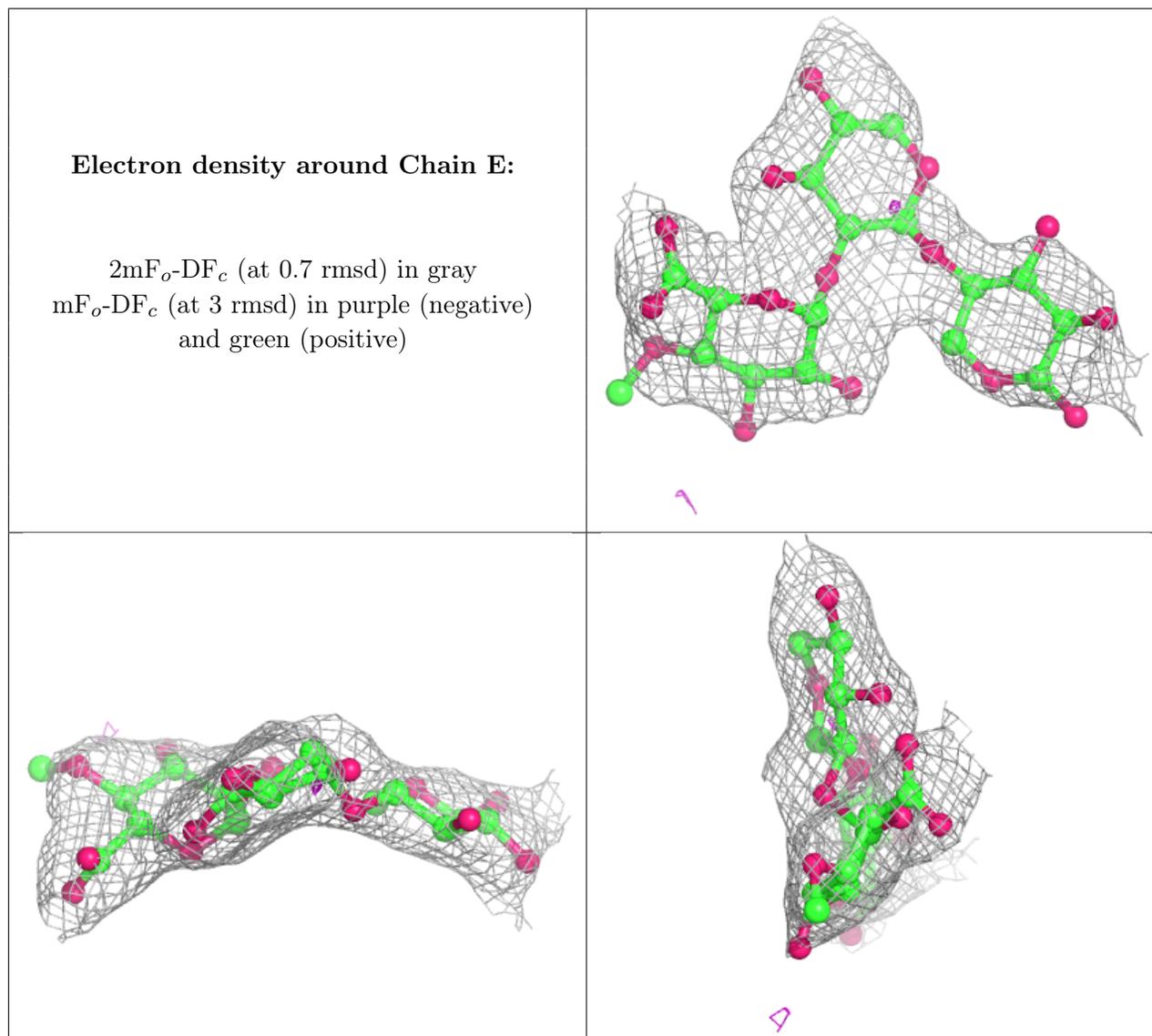
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	XYP	H	1	10/10	0.39	0.41	118,118,118,119	0
2	XYP	H	2	9/10	0.55	0.38	115,116,117,117	0

Continued on next page...

Continued from previous page...

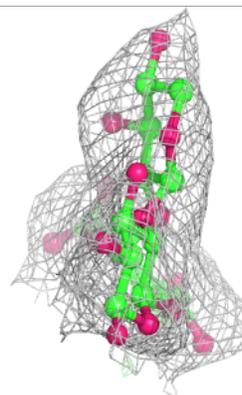
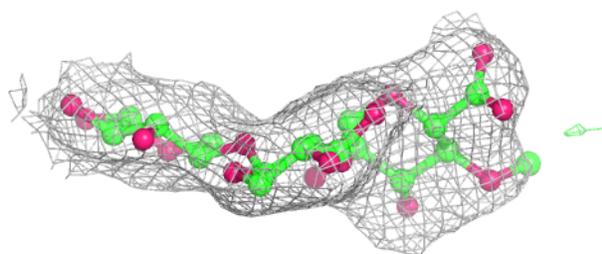
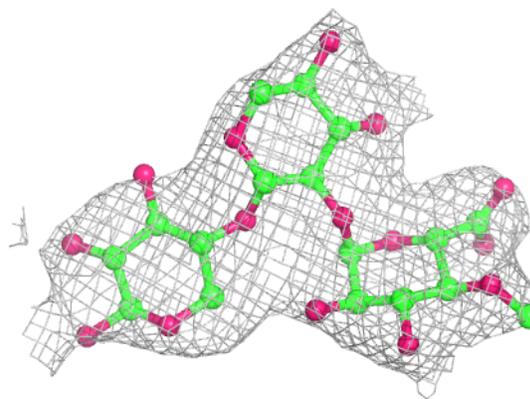
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	XYP	G	2	9/10	0.79	0.17	88,90,91,91	0
2	XYP	E	2	9/10	0.80	0.23	92,93,94,95	0
2	XYP	F	1	10/10	0.83	0.20	80,82,83,83	0
2	GCV	G	3	13/14	0.85	0.19	86,88,89,90	0
2	XYP	E	1	10/10	0.87	0.35	92,94,94,95	0
2	GCV	H	3	13/14	0.87	0.28	112,114,115,115	0
2	GCV	E	3	13/14	0.88	0.30	91,94,95,95	0
2	XYP	G	1	9/10	0.90	0.17	91,92,92,93	0
2	XYP	F	2	9/10	0.94	0.10	76,77,78,78	0
2	GCV	F	3	13/14	0.94	0.17	73,76,78,79	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.

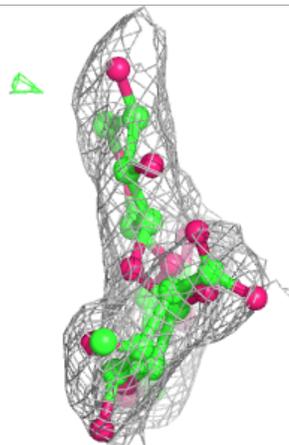
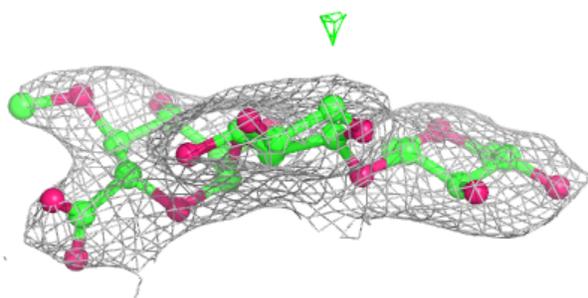
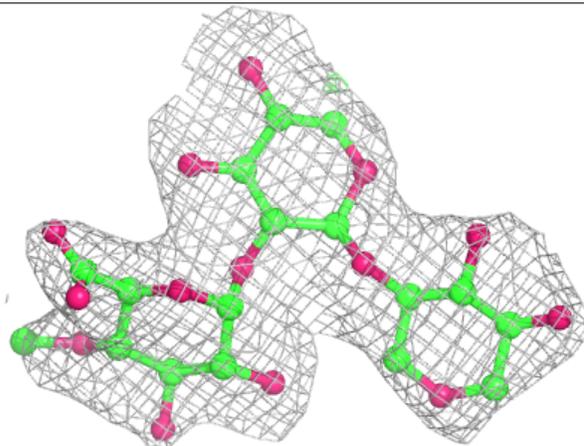


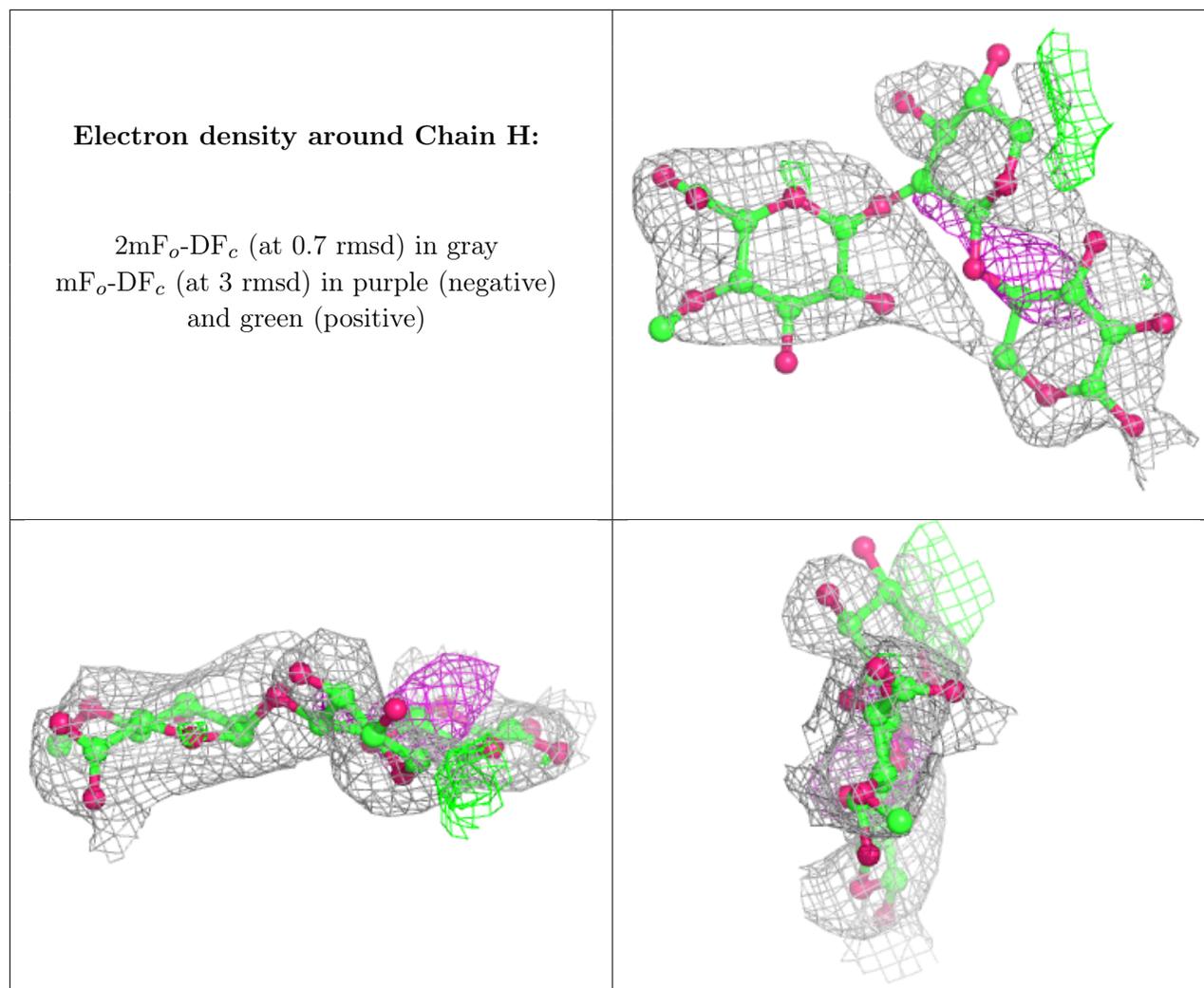
Electron density around Chain F:

$2mF_o-DF_c$ (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.