



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 5, 2023 – 09:54 pm GMT

PDB ID : 8CB1
Title : Crystal structure of human lysosomal acid-alpha-glucosidase, GAA, in complex with N-PNT-DNM 15
Authors : Sulzenbacher, G.; Roig-Zamboni, V.; Overkleeft, H.; Artola, M.
Deposited on : 2023-01-25
Resolution : 1.75 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
buster-report : 1.1.7 (2018)
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

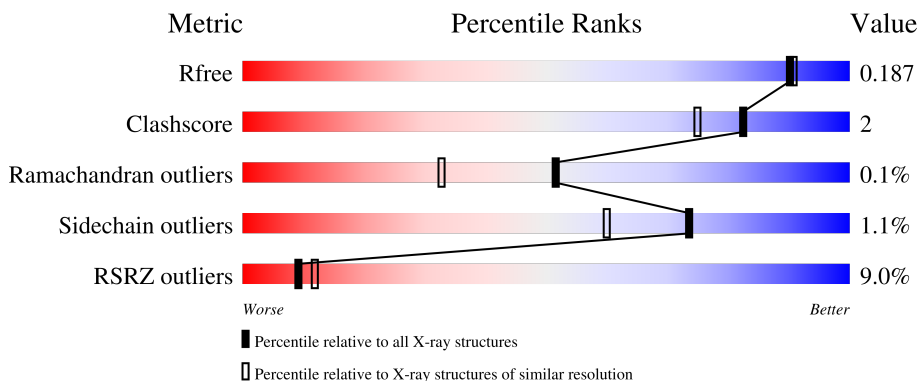
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.75 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.



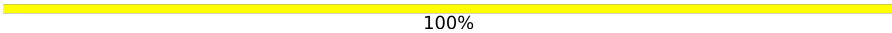
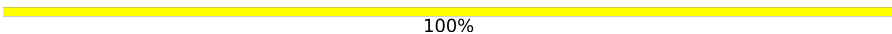
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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	G	123	 11% 88% 1% 1%
2	A	749	 8% 92% 1% 1%
3	B	4	 50% 50%
4	C	2	 100%
4	D	2	 100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	3	 100%
6	F	3	 100%

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
3	NAG	B	2	-	-	-	X
3	BMA	B	3	-	-	-	X
3	FUC	B	4	-	-	-	X
4	NAG	D	2	-	-	-	X
6	NAG	F	2	-	-	-	X
6	FUC	F	3	-	-	-	X

2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 7846 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 Lysosomal alpha-glucosidase (76 kDa).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	G	109	894	565	149	170	10	0	3	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	199	ARG	HIS	conflict	UNP P10253

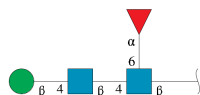
- Molecule 2 is a protein called Lysosomal alpha-glucosidase (70 kDa).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	737	5855	3769	981	1081	24	0	13	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	223	HIS	ARG	conflict	UNP P10253
A	780	ILE	VAL	conflict	UNP P10253

- Molecule 3 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



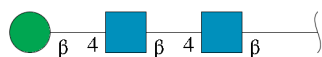
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	B	4	49	28	2	19	0	0	0

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



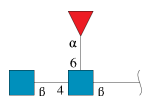
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	C	2	28	16	2	10	0	0	0
4	D	2	28	16	2	10	0	0	0

- Molecule 5 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	E	3	39	22	2	15	0	0	0

- Molecule 6 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
6	F	3	38	22	2	14	0	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	G	2	Total	Cl	0	0
			2	2		
7	A	8	Total	Cl	0	0
			8	8		

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



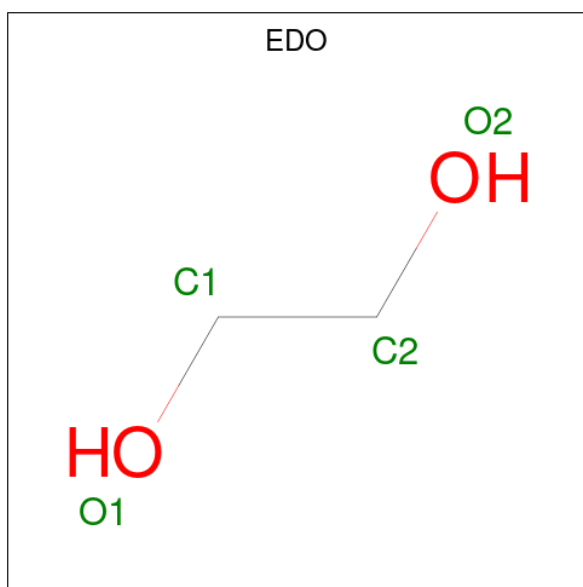
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	G	1	Total	C	O	0	0
			6	3	3		
8	A	1	Total	C	O	0	0
			6	3	3		
8	A	1	Total	C	O	0	0
			6	3	3		
8	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 9 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	G	1	Total	C	O	0	0
			7	4	3		
9	A	1	Total	C	O	0	0
			7	4	3		
9	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 10 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



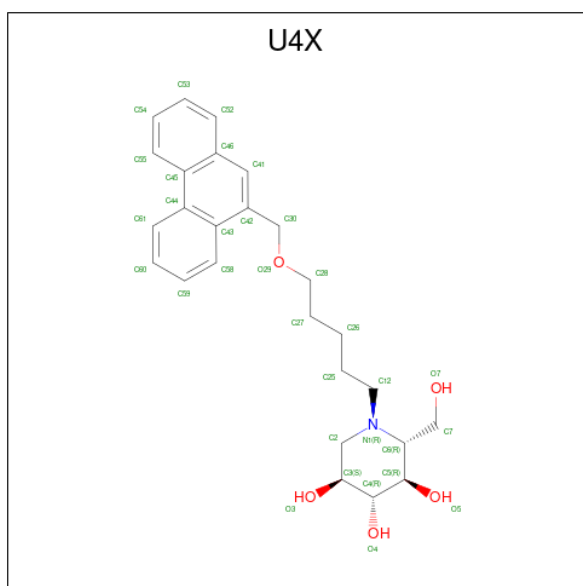
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	G	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

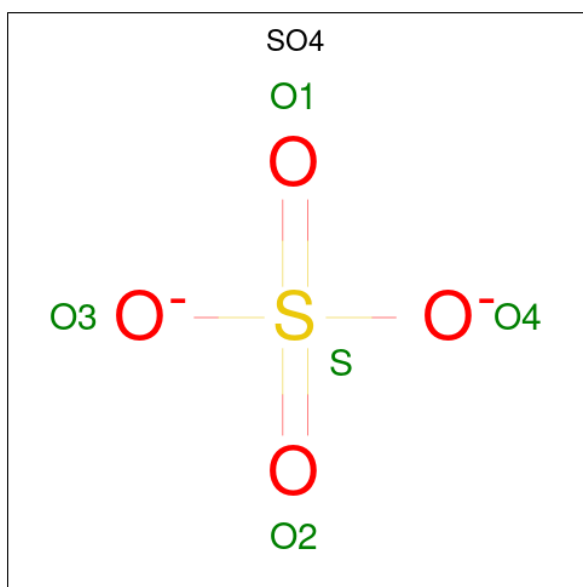
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	G	1	Total	C	O	0	0
			4	2	2		
10	A	1	Total	C	O	0	0
			4	2	2		
10	A	1	Total	C	O	0	0
			4	2	2		
10	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 11 is (2R,3R,4R,5S)-2-(hydroxymethyl)-1-[5-(phenanthren-9-ylmethoxy)pentyl]piperidine-3,4,5-triol (three-letter code: U4X) (formula: C₂₆H₃₃NO₅) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	A	1	Total	C	N	O	0	0
			32	26	1	5		

- Molecule 12 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	A	1	Total O S 5 4 1	0	0
12	A	1	Total O S 5 4 1	0	0
12	A	1	Total O S 5 4 1	0	0
12	A	1	Total O S 5 4 1	0	0

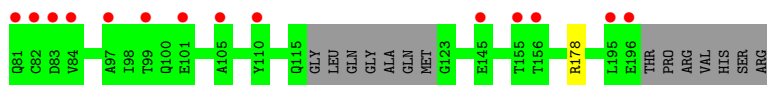
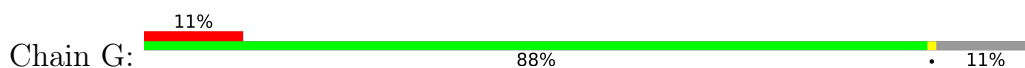
- Molecule 13 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	G	132	Total O 132 132	0	0
13	A	653	Total O 653 653	0	0

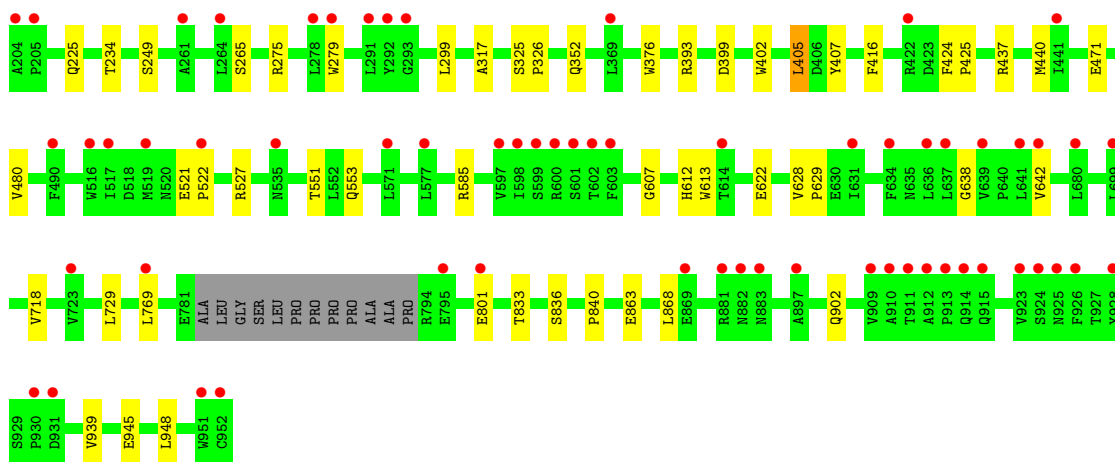
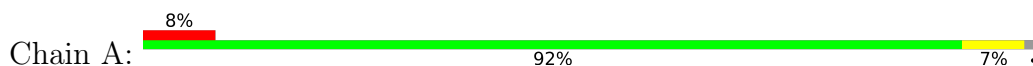
3 Residue-property plots [i](#)

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: Lysosomal alpha-glucosidase (76 kDa)



- Molecule 2: Lysosomal alpha-glucosidase (70 kDa)



- Molecule 3: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose




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



MAG1
MAG2


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

Chain D:  100%MAG1
MAG2

- Molecule 5: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:  100%MAG1
MAG2
BMA3

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:  100%MAG1
MAG2
FUC3

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	96.75Å 102.60Å 129.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.73 – 1.75 47.68 – 1.75	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.73-1.75) 100.0 (47.68-1.75)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.61 (at 1.75Å)	Xtrriage
Refinement program	REFMAC 7.1.015	Depositor
R, R_{free}	0.142 , 0.175 0.155 , 0.187	Depositor DCC
R_{free} test set	6501 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	26.8	Xtrriage
Anisotropy	0.063	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 58.8	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	0.97	EDS
Total number of atoms	7846	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.87% 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: GOL, CSO, BMA, CL, SO4, NAG, PGE, FUC, U4X, EDO

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	G	0.66	0/924	0.87	1/1255 (0.1%)
2	A	0.68	0/6055	0.80	1/8277 (0.0%)
All	All	0.68	0/6979	0.81	2/9532 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	527	ARG	NE-CZ-NH1	-5.21	117.70	120.30
1	G	178	ARG	NE-CZ-NH1	5.08	122.84	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	G	894	0	857	0	0
2	A	5855	0	5690	31	0
3	B	49	0	43	0	0
4	C	28	0	25	0	0
4	D	28	0	25	0	0
5	E	39	0	34	0	0
6	F	38	0	34	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	8	0	0	1	0
7	G	2	0	0	0	0
8	A	18	0	24	1	0
8	G	6	0	8	0	0
9	A	17	0	23	3	0
9	G	7	0	9	0	0
10	A	12	0	18	0	0
10	G	8	0	12	0	0
11	A	32	0	0	0	0
12	A	20	0	0	0	0
13	A	653	0	0	10	0
13	G	132	0	0	0	0
All	All	7846	0	6802	34	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 2.

All (34) 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)
2:A:551:THR:OG1	2:A:553[B]:GLN:OE1	2.03	0.76
2:A:863:GLU:OE1	13:A:1101:HOH:O	2.05	0.75
2:A:585:ARG:NH2	13:A:1106:HOH:O	2.33	0.62
7:A:1006:CL:CL	13:A:1683:HOH:O	2.53	0.61
9:A:1017:PGE:H3	13:A:1180:HOH:O	2.04	0.56
2:A:945:GLU:OE1	9:A:1017:PGE:H32	2.06	0.55
2:A:622:GLU:HG2	13:A:1188:HOH:O	2.06	0.55
2:A:553[B]:GLN:OE1	13:A:1102:HOH:O	2.19	0.53
2:A:437:ARG:CZ	8:A:1016:GOL:H2	2.42	0.50
2:A:405:LEU:H	2:A:405:LEU:HD23	1.78	0.49
2:A:521:GLU:N	2:A:522:PRO:HA	2.29	0.47
2:A:376:TRP:NE1	2:A:405:LEU:HD21	2.30	0.47
2:A:405:LEU:H	2:A:405:LEU:CD2	2.27	0.46
2:A:948:LEU:C	2:A:948:LEU:HD23	2.36	0.45
2:A:718[A]:VAL:HG12	2:A:868:LEU:CD1	2.47	0.45
2:A:393[A]:ARG:NH2	13:A:1109:HOH:O	2.49	0.45
2:A:471:GLU:HB2	13:A:1530:HOH:O	2.17	0.45
2:A:234:THR:HA	2:A:249:SER:O	2.17	0.44
2:A:275:ARG:NH1	13:A:1117:HOH:O	2.51	0.44
9:A:1017:PGE:C3	13:A:1180:HOH:O	2.64	0.43
2:A:279:TRP:CE3	2:A:317:ALA:HB2	2.54	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:325:SER:N	2:A:326:PRO:HA	2.34	0.43
2:A:299:LEU:C	2:A:299:LEU:HD23	2.40	0.42
2:A:628:VAL:HB	2:A:629:PRO:HD3	2.02	0.42
2:A:407:TYR:O	2:A:416:PHE:HA	2.19	0.42
2:A:612:HIS:O	2:A:642:VAL:HA	2.19	0.42
2:A:902[A]:GLN:HA	2:A:939:VAL:O	2.19	0.41
2:A:424:PHE:HB3	2:A:425:PRO:HD3	2.03	0.41
2:A:607:GLY:HA3	2:A:638:GLY:O	2.21	0.41
2:A:769:LEU:HD12	2:A:801:GLU:HA	2.03	0.41
2:A:833:THR:HA	2:A:836:SER:OG	2.21	0.41
2:A:902[B]:GLN:HA	2:A:939:VAL:O	2.21	0.40
2:A:840:PRO:HB3	2:A:902[B]:GLN:HG2	2.04	0.40

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	G	108/123 (88%)	106 (98%)	2 (2%)	0	100	100
2	A	745/749 (100%)	720 (97%)	24 (3%)	1 (0%)	51	33
All	All	853/872 (98%)	826 (97%)	26 (3%)	1 (0%)	51	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	480	VAL

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	G	103/111 (93%)	103 (100%)	0	100	100
2	A	635/630 (101%)	627 (99%)	8 (1%)	69	54
All	All	738/741 (100%)	730 (99%)	8 (1%)	73	60

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	225	GLN
2	A	265	SER
2	A	352	GLN
2	A	399	ASP
2	A	405	LEU
2	A	440	MET
2	A	613	TRP
2	A	729	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	584	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](#)

1 non-standard protein/DNA/RNA residue is 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	CSO	A	938	2	3,6,7	0.80	0	0,6,8	-	-

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	CSO	A	938	2	-	0/1/5/7	-

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.

5.5 Carbohydrates [i](#)

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
3	NAG	B	1	1,3	14,14,15	0.61	0	17,19,21	1.24	1 (5%)
3	NAG	B	2	3	14,14,15	0.60	0	17,19,21	0.88	0
3	BMA	B	3	3	11,11,12	0.60	0	15,15,17	1.32	4 (26%)
3	FUC	B	4	3	10,10,11	0.54	0	14,14,16	1.04	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	NAG	C	1	4,2	14,14,15	0.81	0	17,19,21	1.41	4 (23%)
4	NAG	C	2	4	14,14,15	0.44	0	17,19,21	1.07	1 (5%)
4	NAG	D	1	4,2	14,14,15	0.74	0	17,19,21	1.34	2 (11%)
4	NAG	D	2	4	14,14,15	0.45	0	17,19,21	1.36	1 (5%)
5	NAG	E	1	5,2	14,14,15	0.97	1 (7%)	17,19,21	1.20	1 (5%)
5	NAG	E	2	5	14,14,15	0.75	0	17,19,21	1.33	2 (11%)
5	BMA	E	3	5	11,11,12	0.59	0	15,15,17	1.76	3 (20%)
6	NAG	F	1	2,6	14,14,15	0.61	0	17,19,21	1.25	2 (11%)
6	NAG	F	2	6	14,14,15	0.31	0	17,19,21	1.27	2 (11%)
6	FUC	F	3	6	10,10,11	0.64	0	14,14,16	1.72	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
3	NAG	B	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	B	2	3	-	0/6/23/26	0/1/1/1
3	BMA	B	3	3	-	1/2/19/22	0/1/1/1
3	FUC	B	4	3	-	-	0/1/1/1
4	NAG	C	1	4,2	-	0/6/23/26	0/1/1/1
4	NAG	C	2	4	-	0/6/23/26	0/1/1/1
4	NAG	D	1	4,2	-	0/6/23/26	0/1/1/1
4	NAG	D	2	4	-	0/6/23/26	0/1/1/1
5	NAG	E	1	5,2	-	0/6/23/26	0/1/1/1
5	NAG	E	2	5	-	0/6/23/26	0/1/1/1
5	BMA	E	3	5	-	2/2/19/22	0/1/1/1
6	NAG	F	1	2,6	-	0/6/23/26	0/1/1/1
6	NAG	F	2	6	-	2/6/23/26	0/1/1/1
6	FUC	F	3	6	-	-	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	1	NAG	C1-C2	2.30	1.55	1.52

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	3	BMA	O5-C5-C6	4.26	113.88	107.20
3	B	1	NAG	O5-C1-C2	-4.13	104.77	111.29
4	D	2	NAG	C1-O5-C5	3.74	117.26	112.19
6	F	3	FUC	O5-C5-C4	3.70	116.15	109.52
6	F	1	NAG	C1-O5-C5	-3.48	107.48	112.19
5	E	3	BMA	C1-O5-C5	3.16	116.47	112.19
4	C	1	NAG	O5-C5-C6	3.15	112.14	107.20
4	D	1	NAG	O5-C1-C2	-3.12	106.37	111.29
6	F	3	FUC	O5-C1-C2	-2.99	106.15	110.77
5	E	1	NAG	C4-C3-C2	-2.90	106.77	111.02
6	F	3	FUC	C3-C4-C5	2.84	114.20	109.77
5	E	2	NAG	O7-C7-N2	-2.63	117.12	121.95
6	F	2	NAG	C3-C4-C5	-2.61	105.58	110.24
4	C	2	NAG	C1-C2-N2	-2.51	106.21	110.49
6	F	2	NAG	O5-C5-C4	-2.47	104.83	110.83
4	C	1	NAG	O4-C4-C3	-2.46	104.67	110.35
6	F	1	NAG	C1-C2-N2	2.31	114.44	110.49
3	B	3	BMA	O5-C5-C6	2.23	110.69	107.20
4	D	1	NAG	C1-O5-C5	-2.22	109.19	112.19
3	B	3	BMA	O2-C2-C3	2.19	114.53	110.14
4	C	1	NAG	C6-C5-C4	-2.18	107.89	113.00
5	E	2	NAG	O6-C6-C5	-2.17	103.83	111.29
3	B	3	BMA	C1-O5-C5	2.14	115.09	112.19
4	C	1	NAG	O5-C1-C2	-2.09	107.98	111.29
5	E	3	BMA	C6-C5-C4	-2.02	108.27	113.00
3	B	3	BMA	C3-C4-C5	2.01	113.83	110.24

There are no chirality outliers.

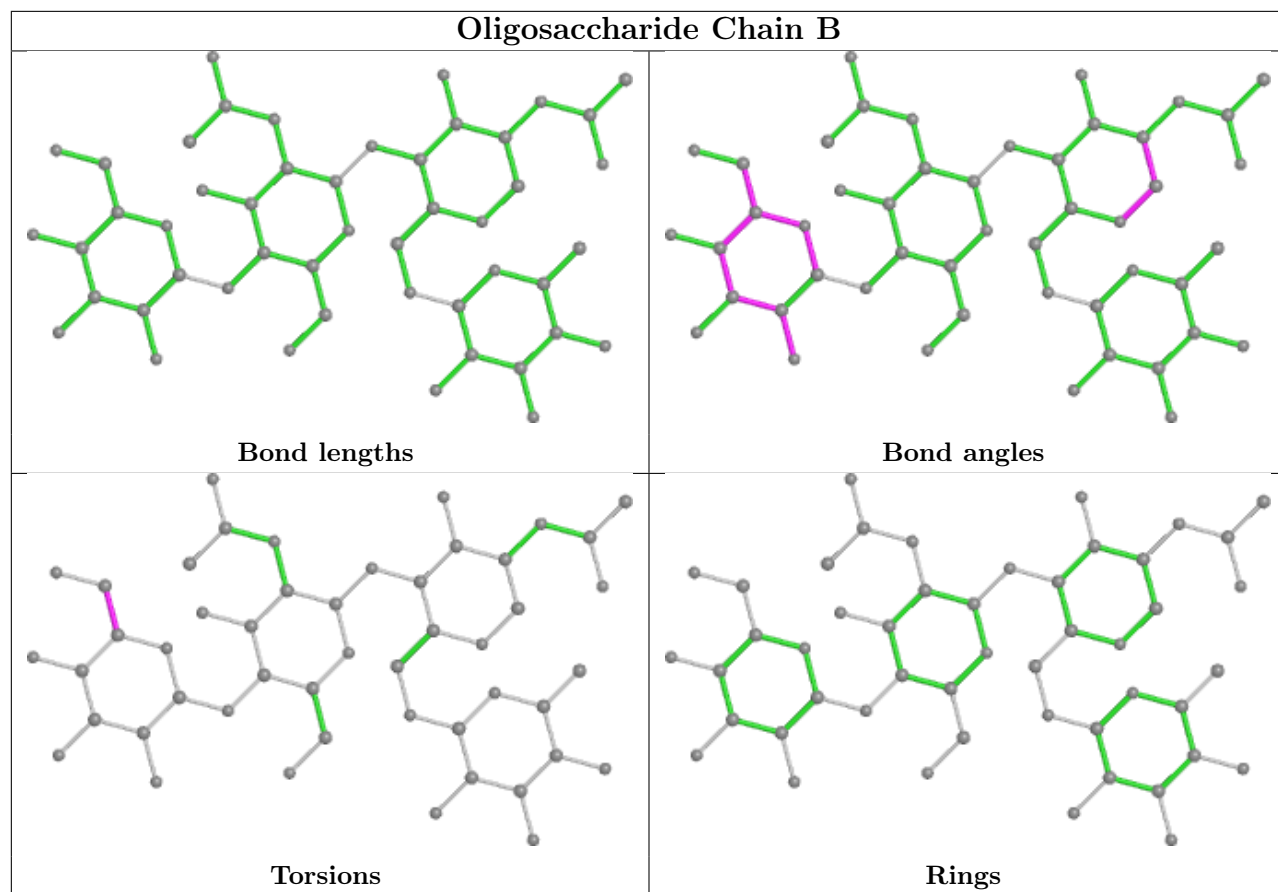
All (5) torsion outliers are listed below:

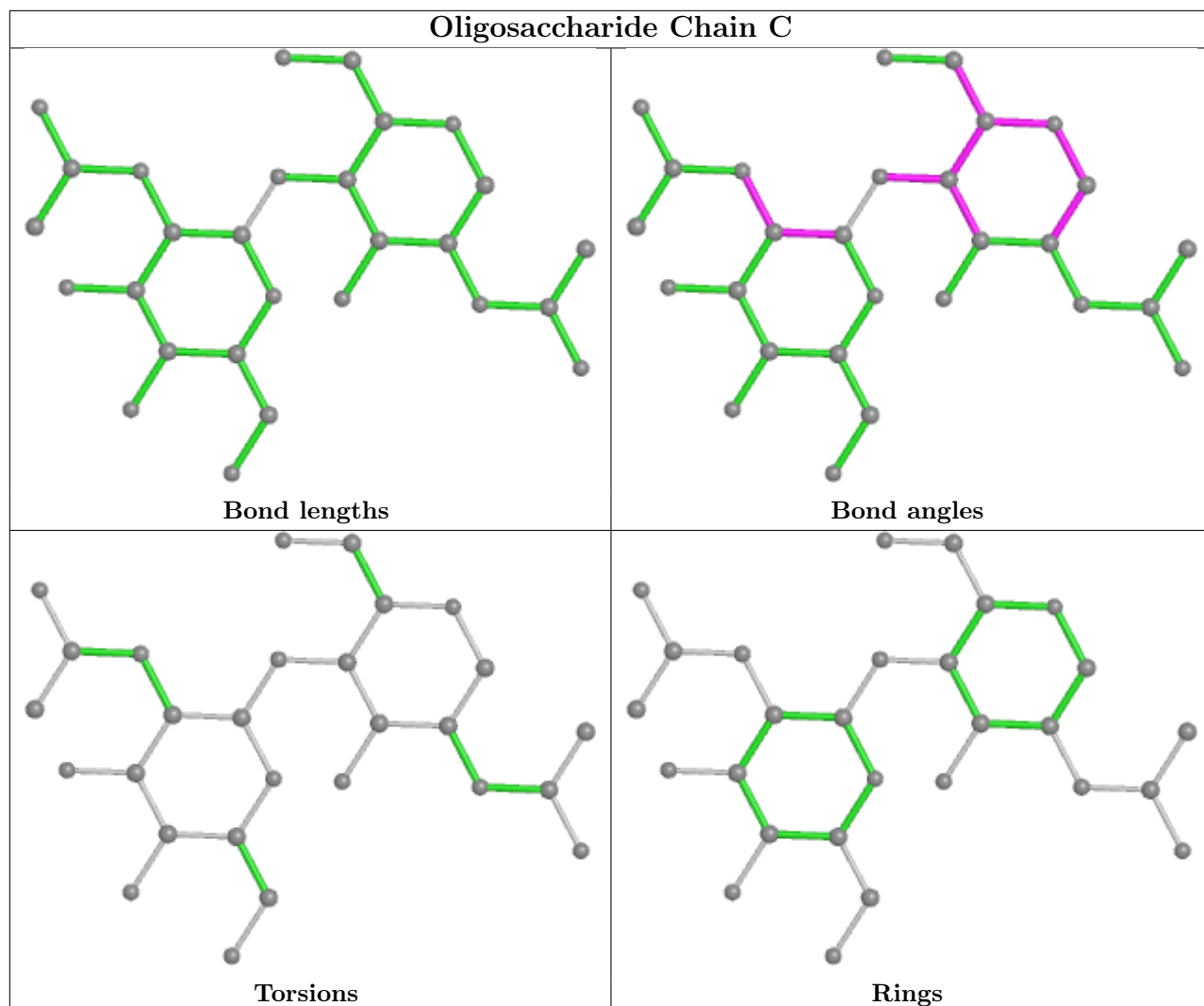
Mol	Chain	Res	Type	Atoms
5	E	3	BMA	O5-C5-C6-O6
6	F	2	NAG	O5-C5-C6-O6
6	F	2	NAG	C4-C5-C6-O6
3	B	3	BMA	O5-C5-C6-O6
5	E	3	BMA	C4-C5-C6-O6

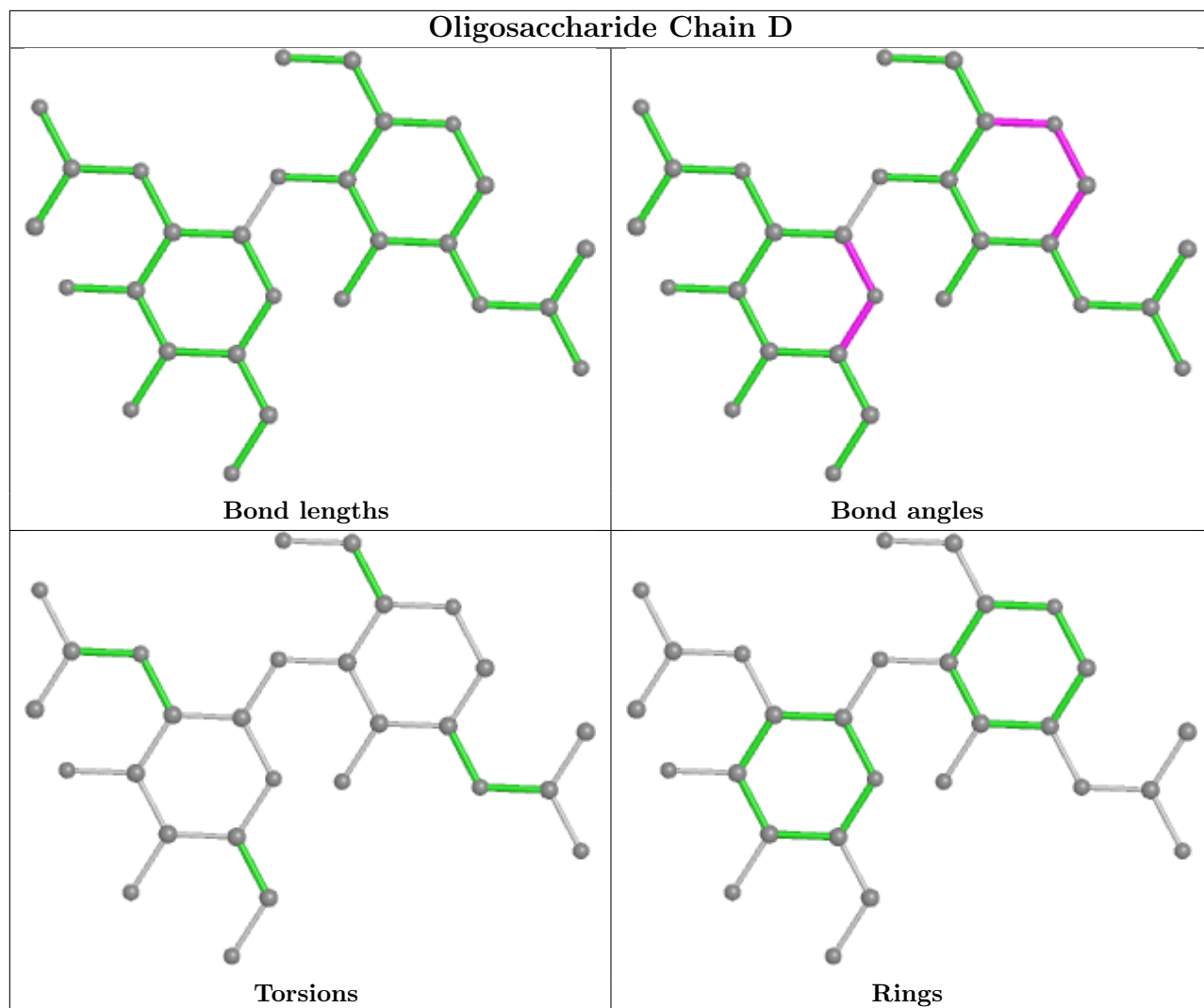
There are no ring outliers.

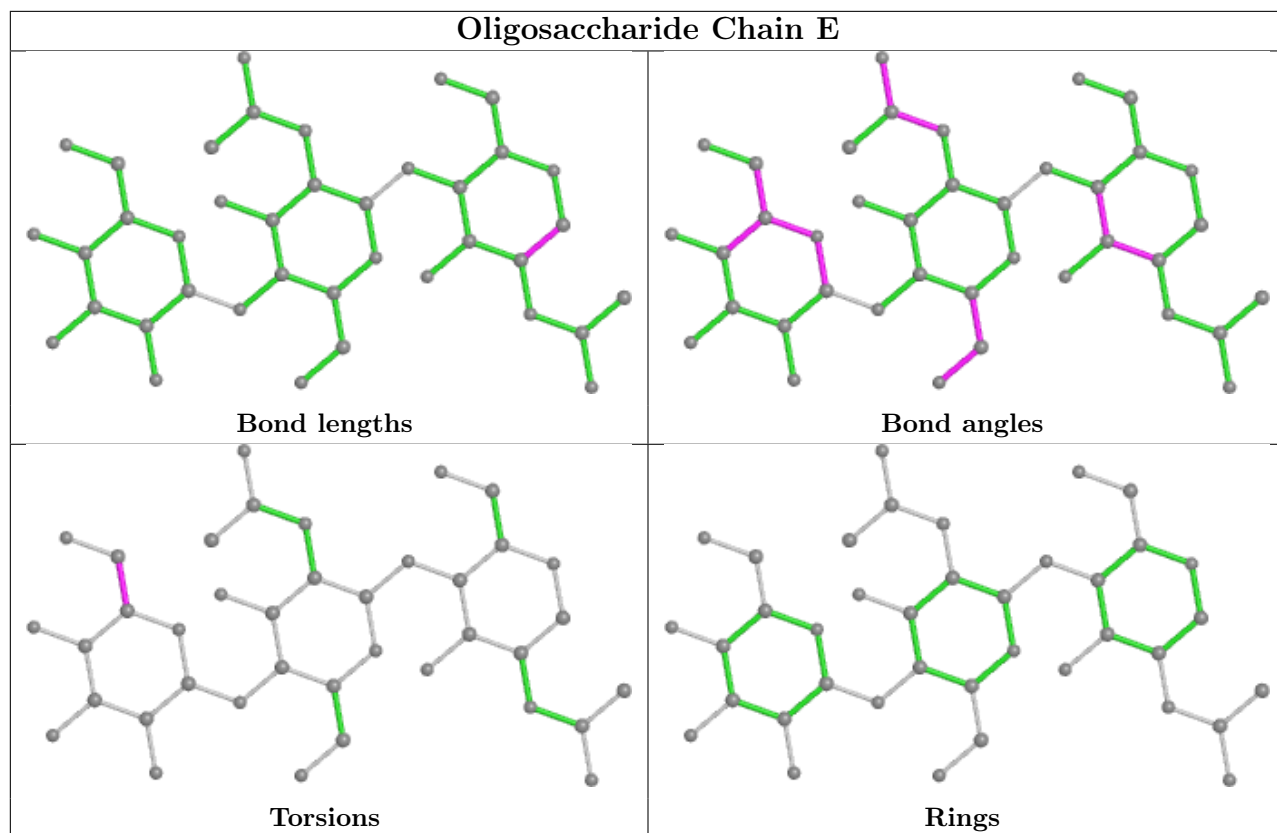
No monomer is involved in short contacts.

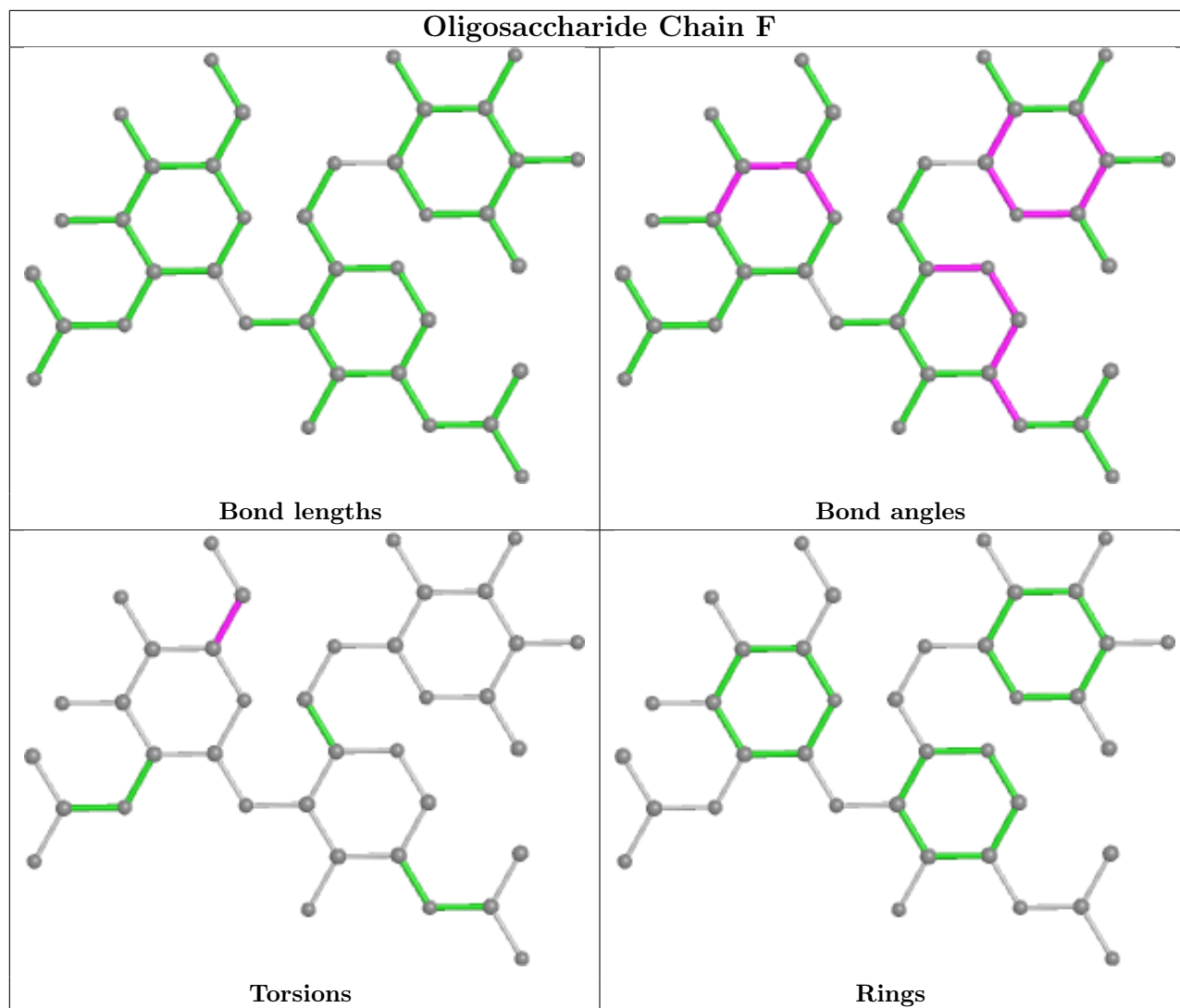
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.











5.6 Ligand geometry [i](#)

Of 27 ligands modelled in this entry, 10 are monoatomic - leaving 17 for Mogul analysis.

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	$\# Z > 2$
9	PGE	G	304	-	6,6,9	0.34	0	5,5,8	0.31	0
10	EDO	G	305	-	3,3,3	0.45	0	2,2,2	0.37	0
8	GOL	A	1015	-	5,5,5	0.31	0	5,5,5	0.42	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
12	SO4	A	1005	-	4,4,4	0.24	0	6,6,6	0.47	0
10	EDO	A	1021	-	3,3,3	0.38	0	2,2,2	0.36	0
8	GOL	A	1016	-	5,5,5	0.23	0	5,5,5	0.66	0
10	EDO	A	1020	-	3,3,3	0.31	0	2,2,2	0.08	0
10	EDO	A	1019	-	3,3,3	0.17	0	2,2,2	0.44	0
12	SO4	A	1002	-	4,4,4	0.31	0	6,6,6	0.27	0
11	U4X	A	1001	-	35,35,35	1.74	6 (17%)	46,48,48	0.93	1 (2%)
10	EDO	G	306	-	3,3,3	0.14	0	2,2,2	0.30	0
8	GOL	G	303	-	5,5,5	0.21	0	5,5,5	0.47	0
9	PGE	A	1018	-	9,9,9	0.25	0	8,8,8	0.25	0
8	GOL	A	1014	-	5,5,5	0.13	0	5,5,5	0.18	0
9	PGE	A	1017	-	6,6,9	0.25	0	5,5,8	0.50	0
12	SO4	A	1003	-	4,4,4	0.27	0	6,6,6	0.25	0
12	SO4	A	1004	-	4,4,4	0.39	0	6,6,6	0.26	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
9	PGE	G	304	-	-	1/4/4/7	-
10	EDO	G	305	-	-	0/1/1/1	-
8	GOL	A	1015	-	-	0/4/4/4	-
10	EDO	A	1021	-	-	0/1/1/1	-
8	GOL	A	1016	-	-	1/4/4/4	-
10	EDO	A	1020	-	-	1/1/1/1	-
10	EDO	A	1019	-	-	0/1/1/1	-
11	U4X	A	1001	-	-	3/12/32/32	0/4/4/4
10	EDO	G	306	-	-	0/1/1/1	-
8	GOL	G	303	-	-	0/4/4/4	-
9	PGE	A	1018	-	-	6/7/7/7	-
8	GOL	A	1014	-	-	0/4/4/4	-
9	PGE	A	1017	-	-	3/4/4/7	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	A	1001	U4X	C43-C44	5.00	1.50	1.42
11	A	1001	U4X	C45-C46	4.88	1.51	1.42
11	A	1001	U4X	C42-C43	4.33	1.52	1.42

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	A	1001	U4X	C45-C44	2.35	1.52	1.45
11	A	1001	U4X	C59-C58	2.22	1.41	1.36
11	A	1001	U4X	C41-C42	2.20	1.42	1.36

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	A	1001	U4X	C43-C44-C45	-2.14	117.10	119.91

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
11	A	1001	U4X	O29-C30-C42-C43
9	A	1018	PGE	O2-C3-C4-O3
9	A	1018	PGE	C4-C3-O2-C2
8	A	1016	GOL	C1-C2-C3-O3
9	A	1017	PGE	O1-C1-C2-O2
11	A	1001	U4X	O29-C30-C42-C41
9	G	304	PGE	O2-C3-C4-O3
9	A	1018	PGE	O3-C5-C6-O4
9	A	1018	PGE	O1-C1-C2-O2
9	A	1018	PGE	C1-C2-O2-C3
9	A	1017	PGE	C4-C3-O2-C2
9	A	1017	PGE	C1-C2-O2-C3
9	A	1018	PGE	C3-C4-O3-C5
10	A	1020	EDO	O1-C1-C2-O2
11	A	1001	U4X	C25-C26-C27-C28

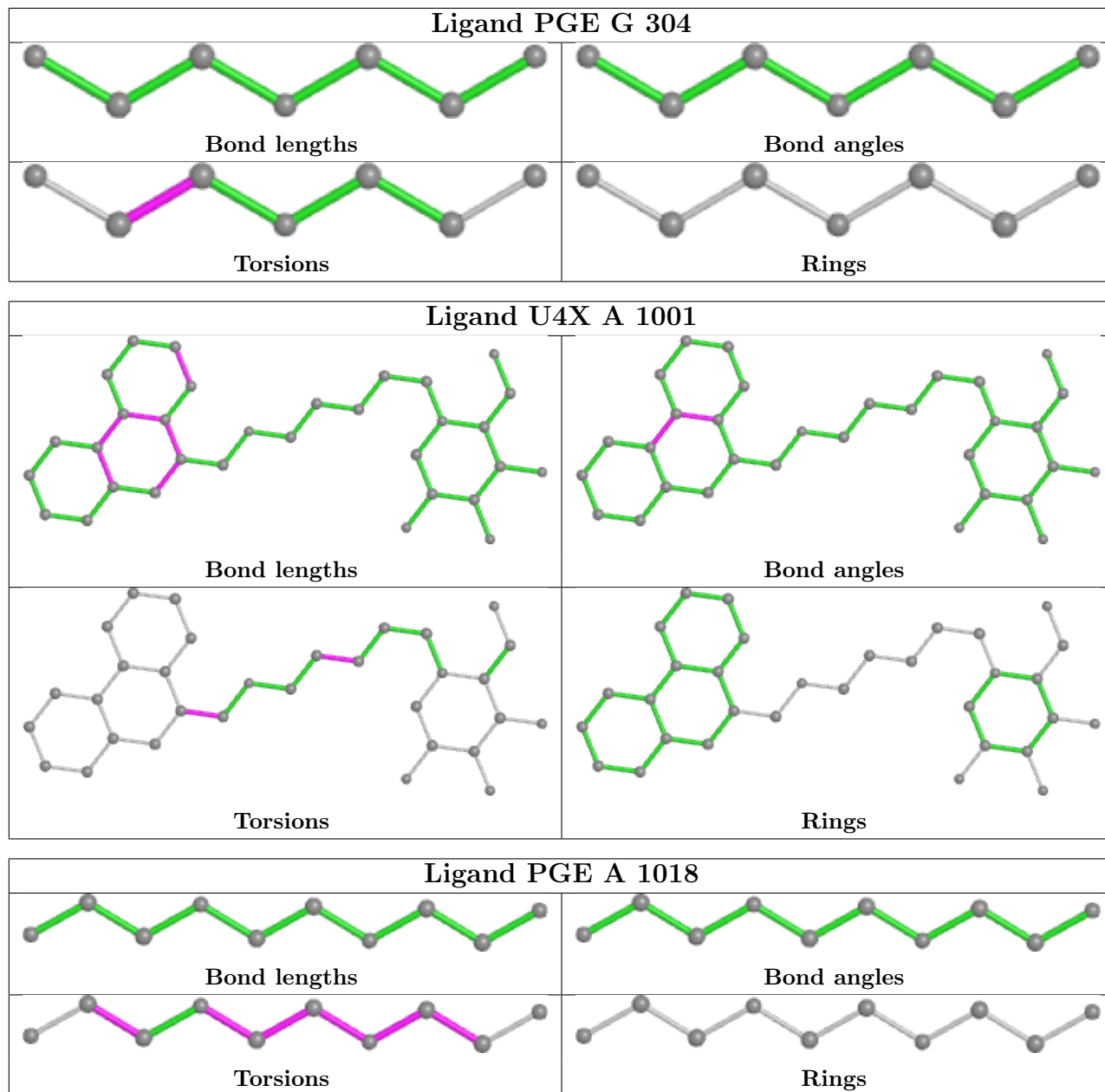
There are no ring outliers.

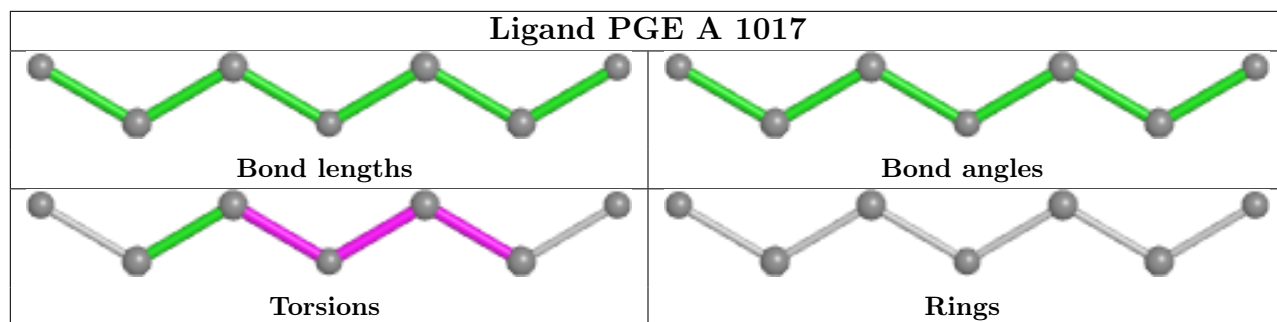
2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	1016	GOL	1	0
9	A	1017	PGE	3	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier.

Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





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

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	G	109/123 (88%)	0.82	14 (12%) 3 5	35, 43, 71, 95	0
2	A	736/749 (98%)	0.66	62 (8%) 11 14	32, 38, 57, 92	0
All	All	845/872 (96%)	0.68	76 (8%) 9 12	32, 39, 59, 95	0

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	952	CYS	7.2
2	A	205	PRO	6.6
2	A	911	THR	5.5
1	G	195	LEU	5.4
2	A	914	GLN	5.1
2	A	912	ALA	4.8
2	A	951	TRP	4.5
1	G	83	ASP	4.3
1	G	84	VAL	4.2
2	A	291	LEU	4.0
2	A	915	GLN	3.8
1	G	81	GLN	3.8
2	A	769	LEU	3.7
2	A	931	ASP	3.7
2	A	204	ALA	3.6
2	A	535	ASN	3.5
2	A	913	PRO	3.5
2	A	422	ARG	3.4
2	A	925	ASN	3.4
2	A	642	VAL	3.3
1	G	105	ALA	3.3
2	A	881	ARG	3.3
2	A	924	SER	3.2
2	A	292	TYR	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	A	882	ASN	3.1
2	A	278	LEU	3.1
2	A	639	VAL	3.1
2	A	930	PRO	3.0
2	A	910	ALA	3.0
2	A	631	ILE	2.9
2	A	909	VAL	2.9
1	G	97	ALA	2.9
1	G	196	GLU	2.9
2	A	801	GLU	2.8
2	A	926	PHE	2.8
2	A	883	ASN	2.8
1	G	101	GLU	2.8
2	A	517	ILE	2.8
2	A	869	GLU	2.7
1	G	99	THR	2.7
2	A	897	ALA	2.7
2	A	641	LEU	2.7
2	A	602	THR	2.7
2	A	279	TRP	2.7
2	A	598	ILE	2.6
2	A	441	ILE	2.6
2	A	261	ALA	2.5
1	G	155	THR	2.5
1	G	156	THR	2.5
2	A	603	PHE	2.5
2	A	577	LEU	2.5
2	A	928	TYR	2.4
2	A	369	LEU	2.4
2	A	636	LEU	2.4
2	A	634	PHE	2.4
2	A	795	GLU	2.4
1	G	110	TYR	2.3
2	A	264	LEU	2.3
2	A	923	VAL	2.2
2	A	571	LEU	2.2
2	A	599	SER	2.2
2	A	680	LEU	2.2
2	A	614	THR	2.2
2	A	723	VAL	2.2
2	A	490	PHE	2.2
1	G	82	CYS	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	A	600	ARG	2.2
2	A	637	LEU	2.1
2	A	519[A]	MET	2.1
1	G	145	GLU	2.1
2	A	522	PRO	2.1
2	A	293	GLY	2.1
2	A	601	SER	2.1
2	A	699	LEU	2.0
2	A	516	TRP	2.0
2	A	597	VAL	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [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	CSO	A	938	7/8	0.85	0.12	46,47,58,59	0

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.

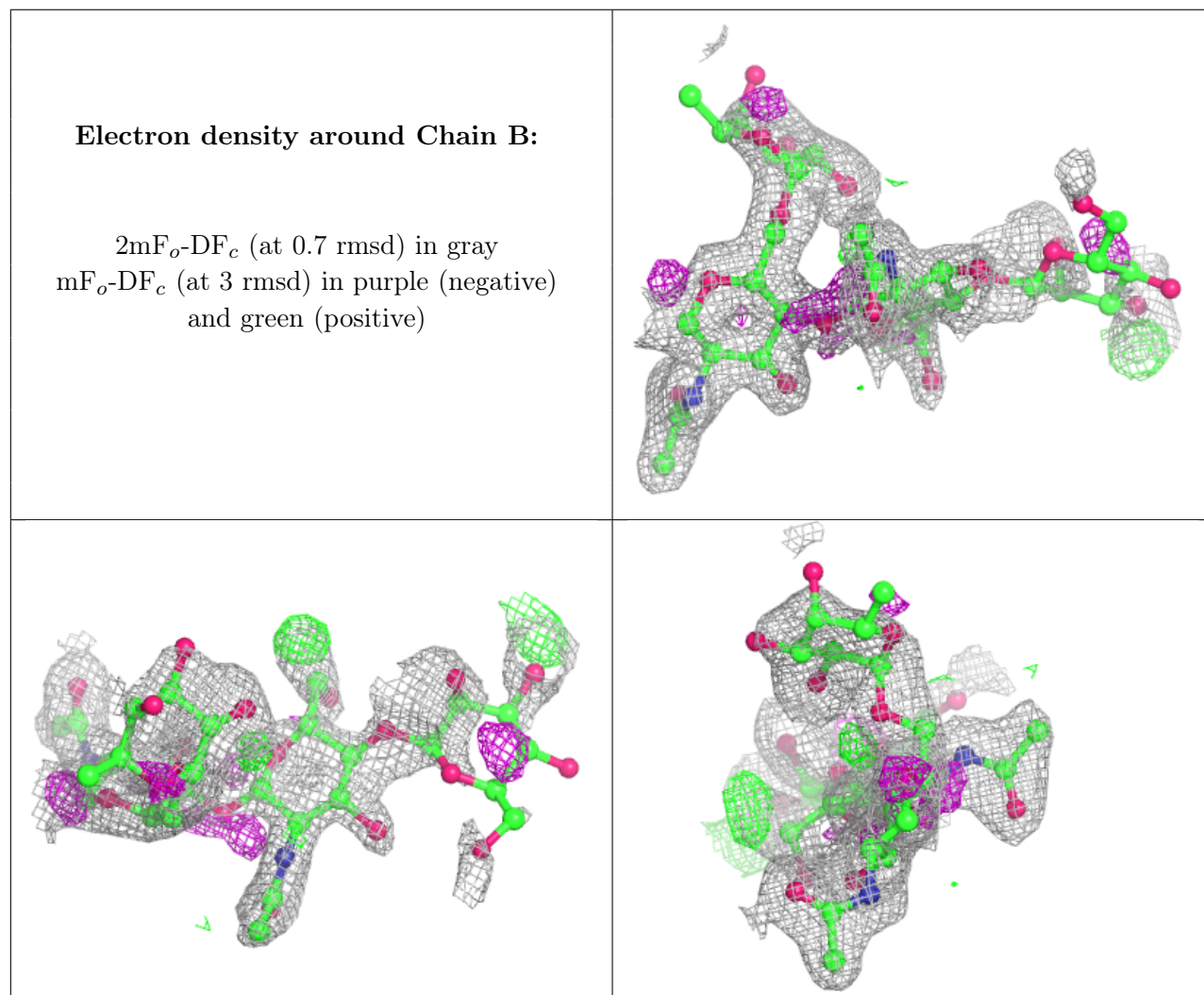
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	BMA	B	3	11/12	0.37	0.64	83,108,123,124	0
6	NAG	F	2	14/15	0.42	0.65	85,119,138,151	0
3	NAG	B	2	14/15	0.65	0.42	61,71,93,105	0
4	NAG	C	2	14/15	0.70	0.35	67,80,96,103	0
4	NAG	D	2	14/15	0.71	0.50	70,93,107,108	0
6	FUC	F	3	10/11	0.75	0.52	86,99,109,117	0
6	NAG	F	1	14/15	0.76	0.37	50,69,80,99	0
3	FUC	B	4	10/11	0.79	0.40	67,82,93,94	0
4	NAG	D	1	14/15	0.79	0.20	42,51,61,70	0
5	BMA	E	3	11/12	0.81	0.34	64,76,91,104	0
3	NAG	B	1	14/15	0.82	0.29	47,52,59,63	0
4	NAG	C	1	14/15	0.85	0.15	37,45,49,54	0

Continued on next page...

Continued from previous page...

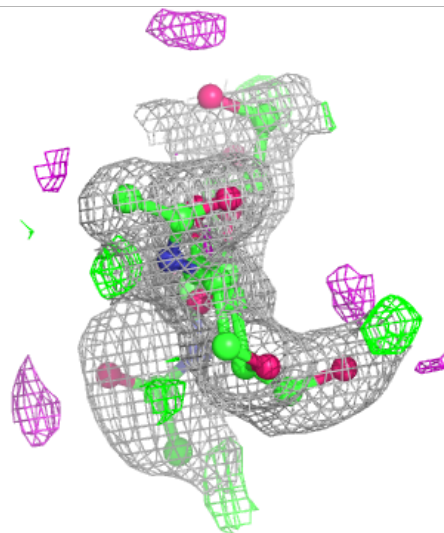
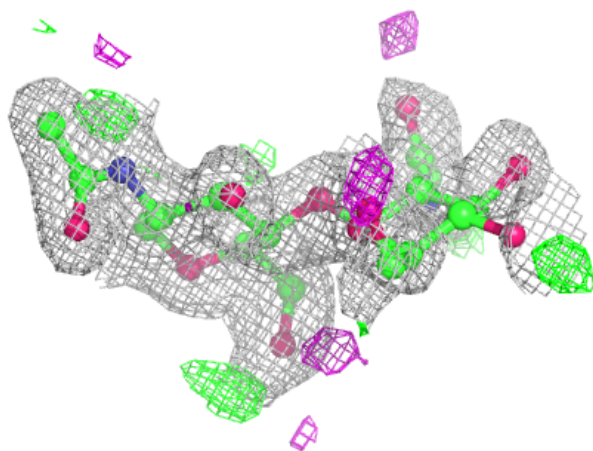
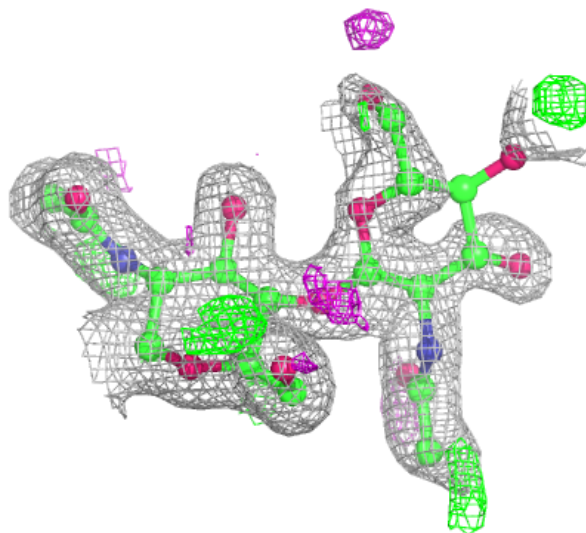
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
5	NAG	E	1	14/15	0.90	0.12	36,41,46,53	0
5	NAG	E	2	14/15	0.90	0.17	42,45,57,57	0

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.



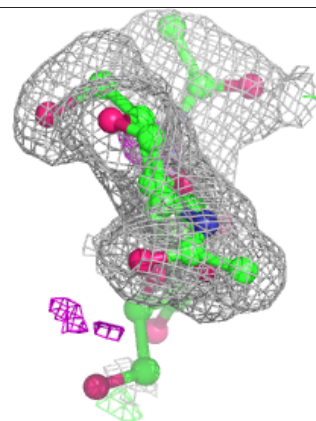
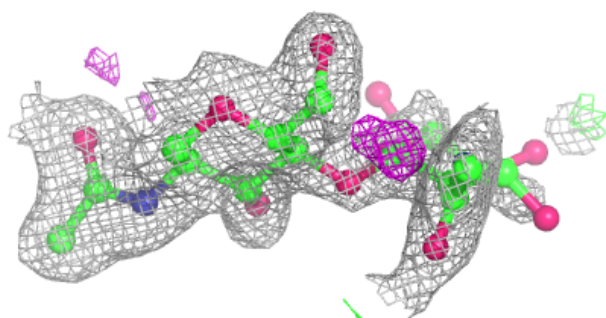
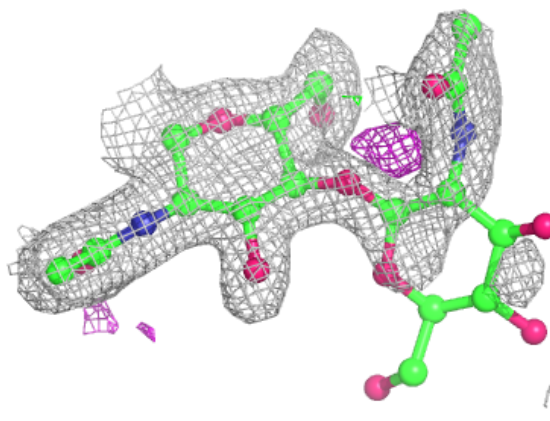
Electron density around Chain C:

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

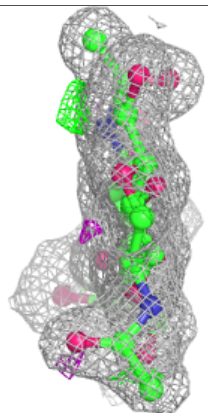
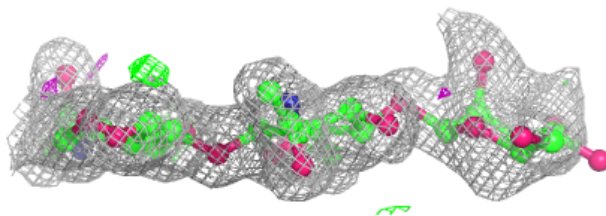
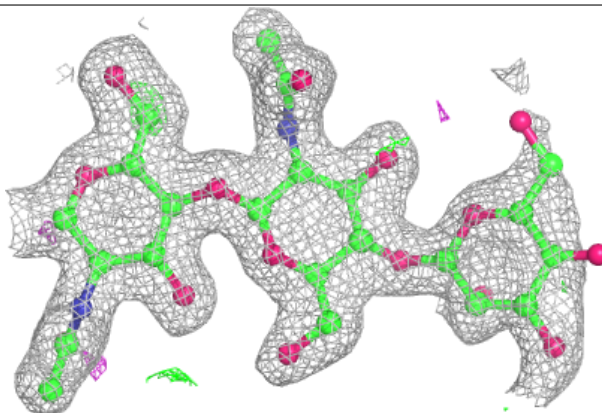


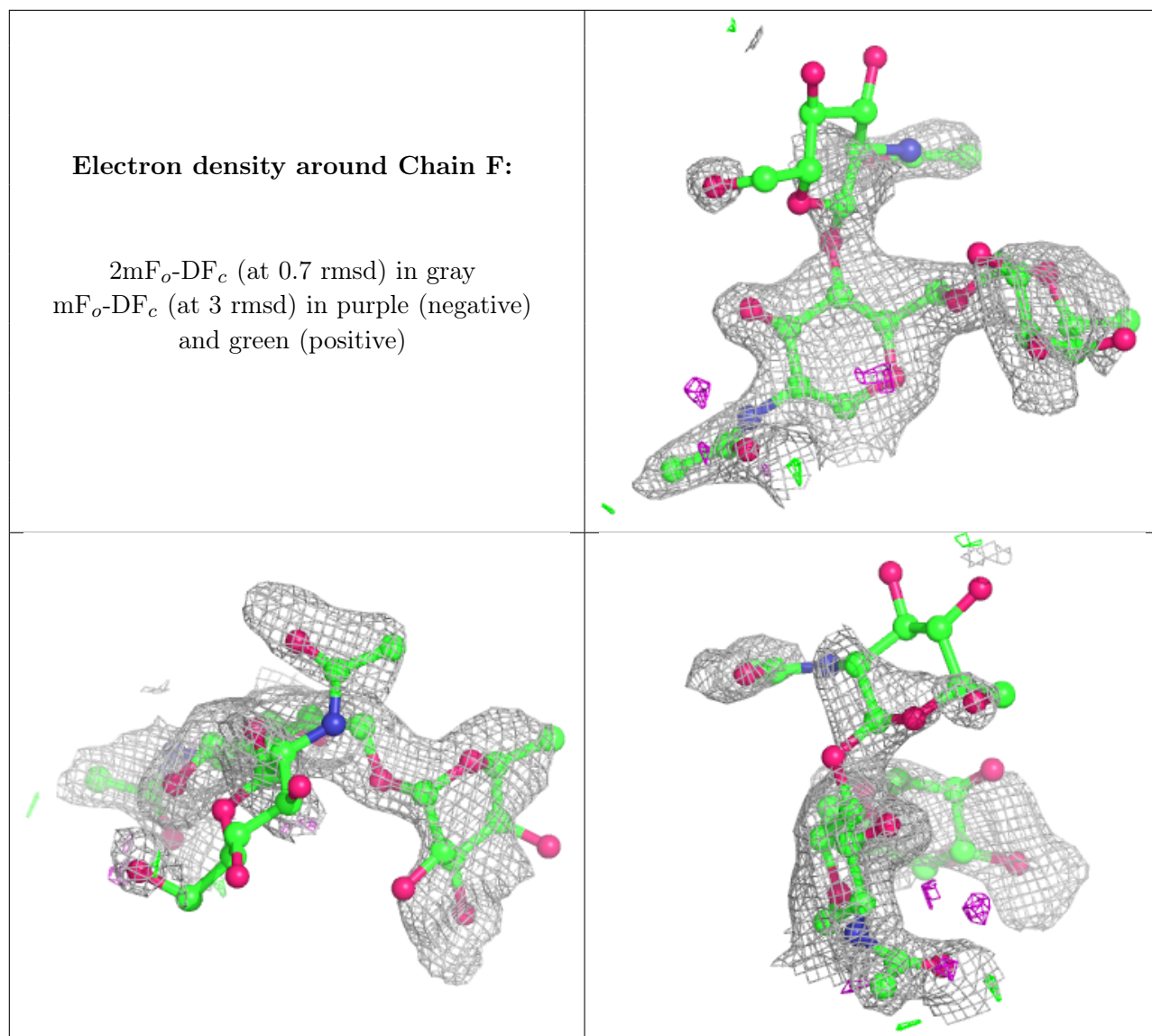
Electron density around Chain D:

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

**Electron density around Chain E:**

$2mF_o-DF_c$ (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
10	EDO	A	1021	4/4	0.62	0.18	64,67,70,86	0
9	PGE	A	1017	7/10	0.63	0.22	57,65,77,78	0
7	CL	A	1011	1/1	0.66	0.33	90,90,90,90	0
10	EDO	A	1020	4/4	0.67	0.29	66,68,68,80	0
8	GOL	A	1015	6/6	0.74	0.26	43,48,51,55	0
7	CL	A	1008	1/1	0.77	0.09	80,80,80,80	0

Continued on next page...

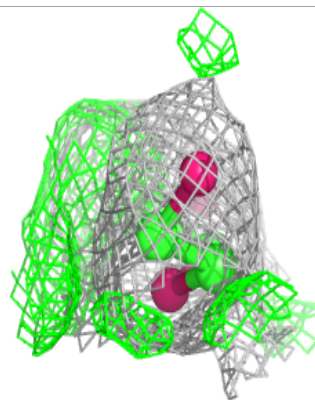
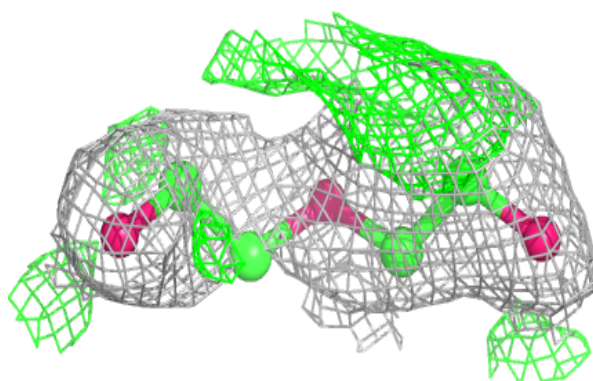
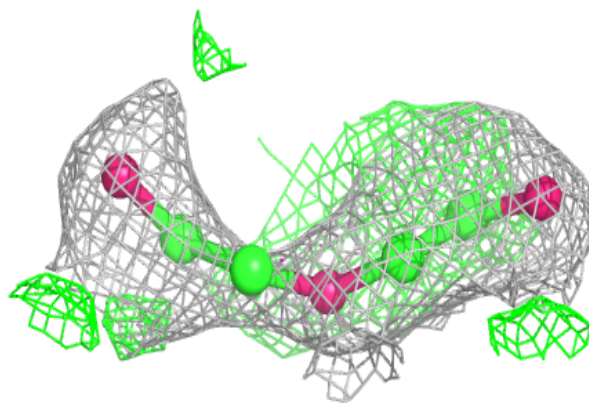
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
9	PGE	G	304	7/10	0.78	0.26	64,70,73,76	0
7	CL	A	1010	1/1	0.78	0.15	84,84,84,84	0
8	GOL	A	1016	6/6	0.84	0.14	56,59,62,67	0
9	PGE	A	1018	10/10	0.85	0.15	62,70,74,77	0
10	EDO	A	1019	4/4	0.85	0.11	63,70,71,73	0
12	SO4	A	1003	5/5	0.85	0.28	65,68,84,99	0
11	U4X	A	1001	32/32	0.87	0.19	32,55,86,102	0
8	GOL	G	303	6/6	0.87	0.12	43,52,52,56	0
7	CL	A	1012	1/1	0.89	0.07	80,80,80,80	0
7	CL	A	1007	1/1	0.89	0.14	79,79,79,79	0
7	CL	G	301	1/1	0.89	0.13	78,78,78,78	0
10	EDO	G	306	4/4	0.90	0.09	63,63,64,66	0
12	SO4	A	1004	5/5	0.91	0.18	62,66,80,83	5
8	GOL	A	1014	6/6	0.92	0.11	43,46,48,49	0
7	CL	G	302	1/1	0.93	0.05	76,76,76,76	0
10	EDO	G	305	4/4	0.94	0.13	48,49,67,68	0
7	CL	A	1006	1/1	0.95	0.11	76,76,76,76	0
12	SO4	A	1005	5/5	0.95	0.11	36,41,47,51	5
7	CL	A	1009	1/1	0.97	0.09	61,61,61,61	0
7	CL	A	1013	1/1	0.97	0.06	71,71,71,71	0
12	SO4	A	1002	5/5	0.98	0.20	48,49,51,53	0

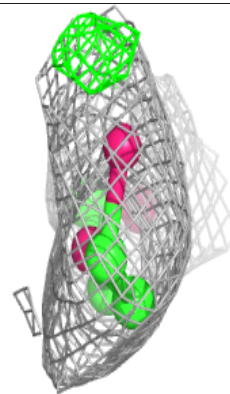
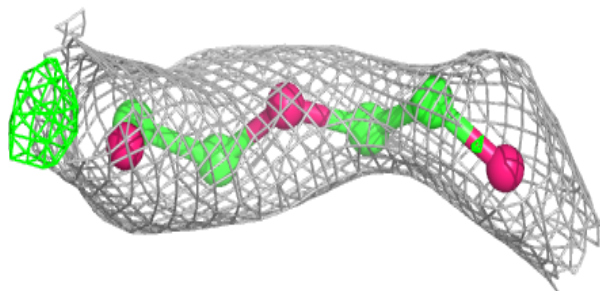
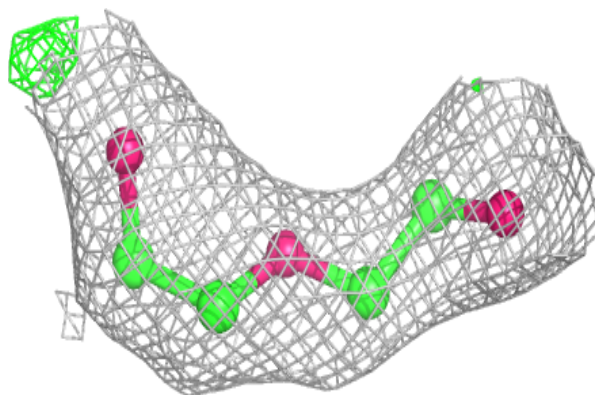
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around PGE A 1017:

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

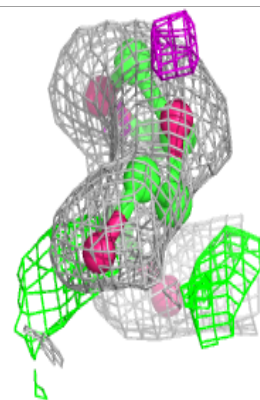
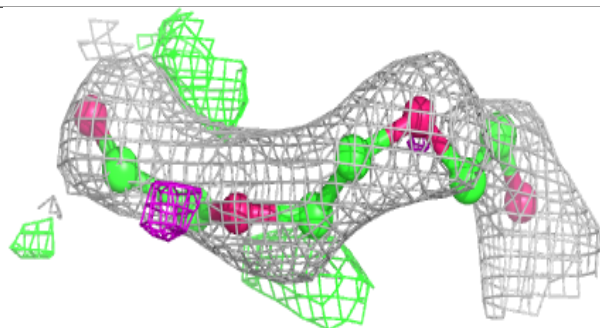
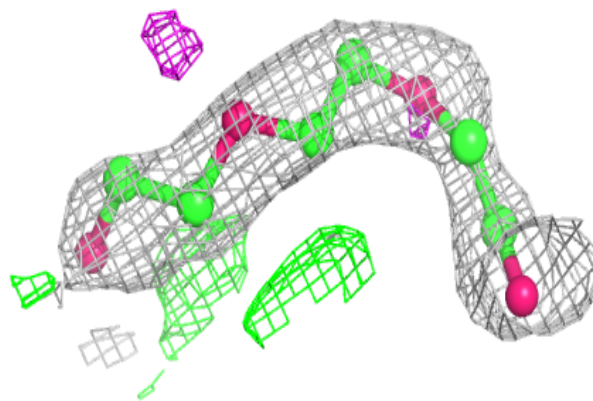
**Electron density around PGE G 304:**

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

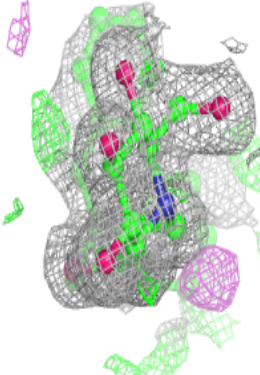
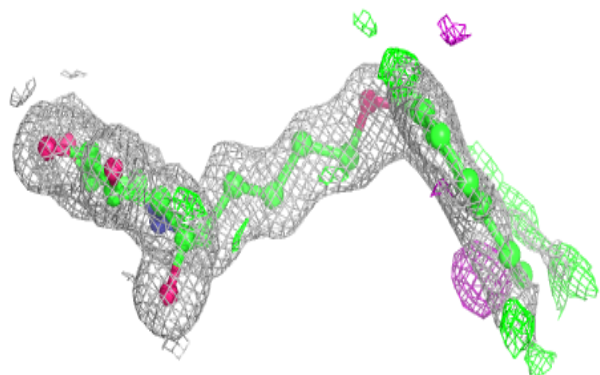
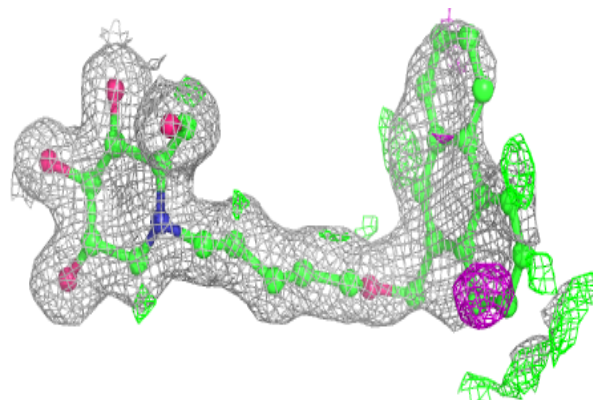


Electron density around PGE A 1018:

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

**Electron density around U4X A 1001:**

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



6.5 Other polymers [i](#)

There are no such residues in this entry.