WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Oct 8, 2023 – 01:40 PM EDT

PDB ID	:	4RFB
Title	:	1.93 Angstrom Crystal Structure of Superantigen-like Protein from Staphylococcus aureus in Complex with Sialyl-Lewis X.
Authors	:	Minasov, G.; Nocadello, S.; Shuvalova, L.; Filippova, E.V.; Halavaty, A.; Dubrovska, I.; Bagnoli, F.; Falugi, F.; Bottomley, M.; Grandi, G.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on	:	2014-09-25
Resolution	:	1.93 Å(reported)

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	:	2.35.1
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

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 1.93 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	Percentile Ranks	Value
Rfree		0.222
Clashscore		4
Ramachandran outliers		0.1%
Sidechain outliers		0.8%
RSRZ outliers		4.3%

Worse .... *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
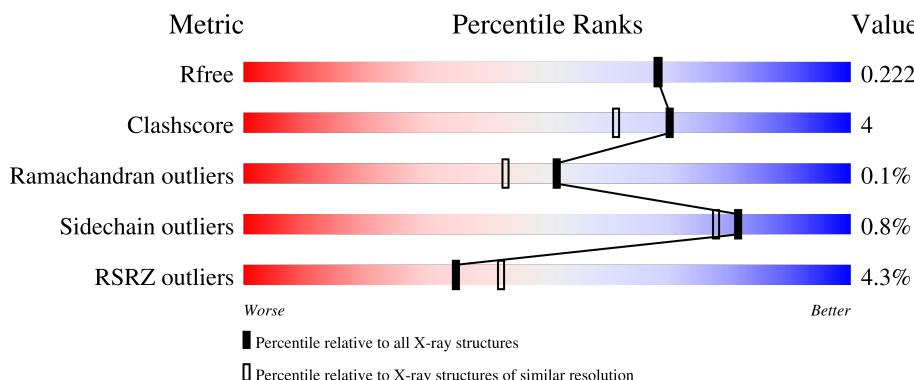
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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% 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	218	2% / 82% / 7% / 11%
1	B	218	4% / 78% / 10% / · / 12%
1	C	218	5% / 82% / 6% / 12%
1	D	218	5% / 81% / 7% / 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	4	75% 25%
2	F	4	75% 25%
2	G	4	25% 50% 25%
2	H	4	25% 75%
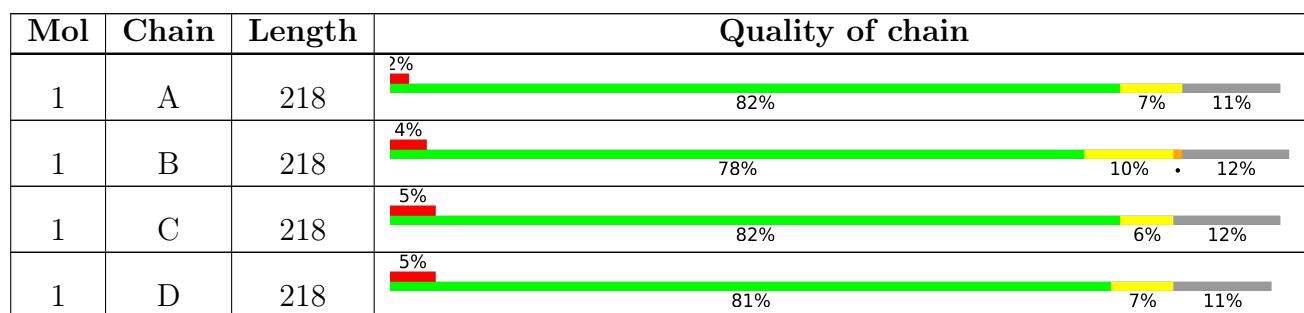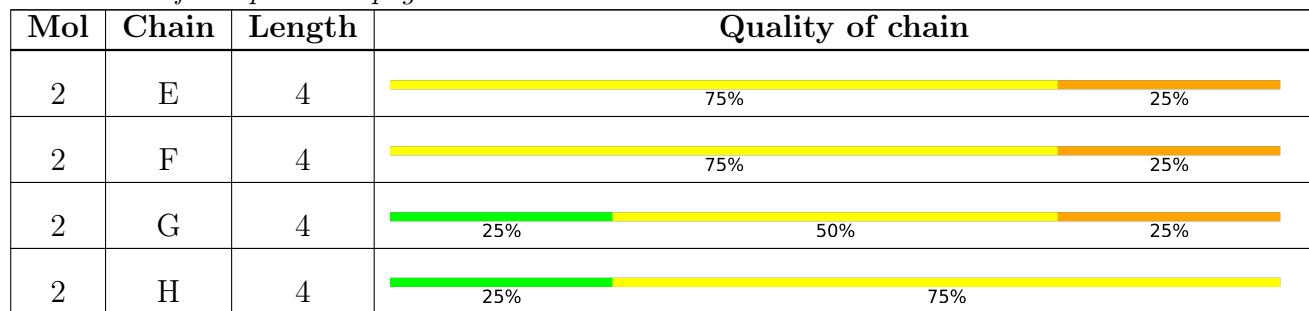
2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 7321 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 Putative uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	194	Total	C	N	O	S	0	7	0
			1668	1061	278	326	3			
1	B	192	Total	C	N	O	S	0	4	0
			1641	1045	277	316	3			
1	C	191	Total	C	N	O	S	0	4	0
			1623	1036	272	311	4			
1	D	193	Total	C	N	O	S	0	4	0
			1638	1043	271	321	3			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-16	MET	-	expression tag	UNP Q2G1S5
A	-15	GLY	-	expression tag	UNP Q2G1S5
A	-14	SER	-	expression tag	UNP Q2G1S5
A	-13	SER	-	expression tag	UNP Q2G1S5
A	-12	HIS	-	expression tag	UNP Q2G1S5
A	-11	HIS	-	expression tag	UNP Q2G1S5
A	-10	HIS	-	expression tag	UNP Q2G1S5
A	-9	HIS	-	expression tag	UNP Q2G1S5
A	-8	HIS	-	expression tag	UNP Q2G1S5
A	-7	HIS	-	expression tag	UNP Q2G1S5
A	-6	GLU	-	expression tag	UNP Q2G1S5
A	-5	ASN	-	expression tag	UNP Q2G1S5
A	-4	LEU	-	expression tag	UNP Q2G1S5
A	-3	TYR	-	expression tag	UNP Q2G1S5
A	-2	PHE	-	expression tag	UNP Q2G1S5
A	-1	GLN	-	expression tag	UNP Q2G1S5
A	0	GLY	-	expression tag	UNP Q2G1S5
B	-16	MET	-	expression tag	UNP Q2G1S5
B	-15	GLY	-	expression tag	UNP Q2G1S5
B	-14	SER	-	expression tag	UNP Q2G1S5
B	-13	SER	-	expression tag	UNP Q2G1S5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-12	HIS	-	expression tag	UNP Q2G1S5
B	-11	HIS	-	expression tag	UNP Q2G1S5
B	-10	HIS	-	expression tag	UNP Q2G1S5
B	-9	HIS	-	expression tag	UNP Q2G1S5
B	-8	HIS	-	expression tag	UNP Q2G1S5
B	-7	HIS	-	expression tag	UNP Q2G1S5
B	-6	GLU	-	expression tag	UNP Q2G1S5
B	-5	ASN	-	expression tag	UNP Q2G1S5
B	-4	LEU	-	expression tag	UNP Q2G1S5
B	-3	TYR	-	expression tag	UNP Q2G1S5
B	-2	PHE	-	expression tag	UNP Q2G1S5
B	-1	GLN	-	expression tag	UNP Q2G1S5
B	0	GLY	-	expression tag	UNP Q2G1S5
C	-16	MET	-	expression tag	UNP Q2G1S5
C	-15	GLY	-	expression tag	UNP Q2G1S5
C	-14	SER	-	expression tag	UNP Q2G1S5
C	-13	SER	-	expression tag	UNP Q2G1S5
C	-12	HIS	-	expression tag	UNP Q2G1S5
C	-11	HIS	-	expression tag	UNP Q2G1S5
C	-10	HIS	-	expression tag	UNP Q2G1S5
C	-9	HIS	-	expression tag	UNP Q2G1S5
C	-8	HIS	-	expression tag	UNP Q2G1S5
C	-7	HIS	-	expression tag	UNP Q2G1S5
C	-6	GLU	-	expression tag	UNP Q2G1S5
C	-5	ASN	-	expression tag	UNP Q2G1S5
C	-4	LEU	-	expression tag	UNP Q2G1S5
C	-3	TYR	-	expression tag	UNP Q2G1S5
C	-2	PHE	-	expression tag	UNP Q2G1S5
C	-1	GLN	-	expression tag	UNP Q2G1S5
C	0	GLY	-	expression tag	UNP Q2G1S5
D	-16	MET	-	expression tag	UNP Q2G1S5
D	-15	GLY	-	expression tag	UNP Q2G1S5
D	-14	SER	-	expression tag	UNP Q2G1S5
D	-13	SER	-	expression tag	UNP Q2G1S5
D	-12	HIS	-	expression tag	UNP Q2G1S5
D	-11	HIS	-	expression tag	UNP Q2G1S5
D	-10	HIS	-	expression tag	UNP Q2G1S5
D	-9	HIS	-	expression tag	UNP Q2G1S5
D	-8	HIS	-	expression tag	UNP Q2G1S5
D	-7	HIS	-	expression tag	UNP Q2G1S5
D	-6	GLU	-	expression tag	UNP Q2G1S5
D	-5	ASN	-	expression tag	UNP Q2G1S5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-4	LEU	-	expression tag	UNP Q2G1S5
D	-3	TYR	-	expression tag	UNP Q2G1S5
D	-2	PHE	-	expression tag	UNP Q2G1S5
D	-1	GLN	-	expression tag	UNP Q2G1S5
D	0	GLY	-	expression tag	UNP Q2G1S5

- Molecule 2 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galacto pyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose.
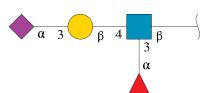
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	4	Total	C	N	O	0	0	0
			56	31	2	23			
2	F	4	Total	C	N	O	0	0	0
			56	31	2	23			
2	G	4	Total	C	N	O	0	0	0
			56	31	2	23			
2	H	4	Total	C	N	O	0	0	0
			56	31	2	23			

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Cl	0	0
			2	2		
3	B	2	Total	Cl	0	0
			2	2		
3	C	1	Total	Cl	0	0
			1	1		
3	D	1	Total	Cl	0	0
			1	1		

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

PEG
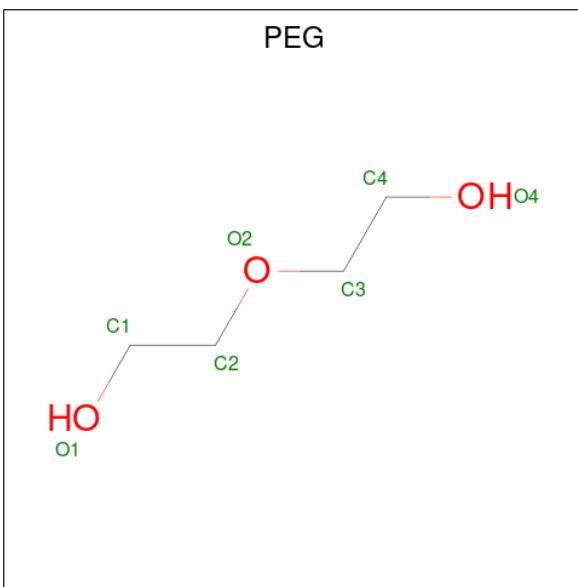
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 7	C 4	O 3	0	0
4	C	1	Total 7	C 4	O 3	0	0
4	C	1	Total 7	C 4	O 3	0	0

• Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).

PGE
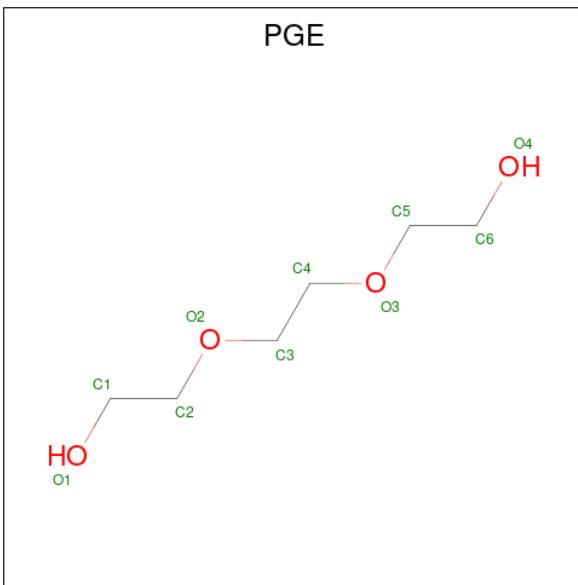
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total 10	C 6	O 4	0	0

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).
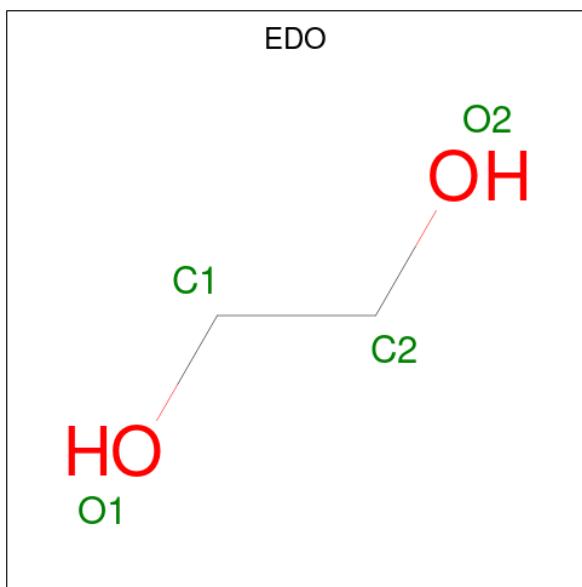
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	D	1	Total 4	C 2	O 2	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	109	Total 112	O 112	0	6
7	B	132	Total 136	O 136	0	6
7	C	114	Total 118	O 118	0	7
7	D	116	Total 120	O 120	0	4

3 Residue-property plots ⓘ

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: Putative uncharacterized protein

Chain A:

- Molecule 1: Putative uncharacterized protein

Chain B:

- Molecule 1: Putative uncharacterized protein

Chain C:

- Molecule 1: Putative uncharacterized protein

Chain D:

N183 N184 R185 Q208 E223 K231

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:

75%　　　　25%

NAG1 GAL2 SIA3 FUC4

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:

75%　　　　25%

NAG1 GAL2 SIA3 FUC4

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:

25%　　　　50%　　　　25%

NAG1 GAL2 SIA3 FUC4

• Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain H:

25%　　　　75%

NAG1 GAL2 SIA3 FUC4

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	42.64Å 108.20Å 167.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.83 – 1.93 28.81 – 1.93	Depositor EDS
% Data completeness (in resolution range)	99.7 (28.83-1.93) 99.7 (28.81-1.93)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	3.41 (at 1.93Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.169 , 0.218 0.176 , 0.222	Depositor DCC
R_{free} test set	2997 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	26.6	Xtriage
Anisotropy	0.298	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 50.2	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7321	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.73% 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: GAL, NAG, SIA, PGE, PEG, EDO, FUC, CL

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	0.58	0/1694	0.75	0/2257
1	B	0.59	0/1666	0.80	0/2216
1	C	0.58	0/1647	0.79	0/2190
1	D	0.55	0/1663	0.72	0/2214
All	All	0.58	0/6670	0.76	0/8877

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	1668	0	1691	14	0
1	B	1641	0	1669	17	0
1	C	1623	0	1665	12	0
1	D	1638	0	1664	9	0
2	E	56	0	49	1	0
2	F	56	0	49	1	0
2	G	56	0	49	1	0
2	H	56	0	49	0	0
3	A	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	2	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	7	0	10	0	0
4	C	14	0	20	0	0
5	B	10	0	14	0	0
6	D	4	0	6	0	0
7	A	112	0	0	4	0
7	B	136	0	0	3	0
7	C	118	0	0	2	0
7	D	120	0	0	1	0
All	All	7321	0	6935	50	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 4.

All (50) 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:140:ILE:HD12	1:C:140:ILE:O	1.90	0.72
1:D:166:PHE:CE2	1:D:170:LYS:HE2	2.24	0.71
1:B:140:ILE:HD12	1:B:141:ASP:N	2.06	0.71
1:D:132:VAL:CG2	1:D:223:GLU:HA	2.24	0.68
1:A:122[A]:LYS:NZ	7:A:507[A]:HOH:O	2.28	0.67
1:C:197[B]:LYS:HB2	1:C:197[B]:LYS:HZ2	1.61	0.66
1:C:39:ASN:OD1	1:C:185[A]:ARG:NH1	2.27	0.65
1:B:84:LEU:HD22	1:B:209:PRO:HB2	1.79	0.65
1:A:222:LYS:HD3	2:G:4:FUC:H62	1.80	0.62
1:C:197[B]:LYS:HB2	1:C:197[B]:LYS:NZ	2.15	0.62
1:B:140:ILE:HD12	1:B:141:ASP:H	1.64	0.60
1:B:61:TYR:CE2	1:B:63:GLN:HA	2.36	0.59
1:D:135[B]:LYS:HE3	1:D:147:SER:HB3	1.86	0.57
1:C:197[A]:LYS:HE3	7:C:470:HOH:O	2.05	0.57
1:B:151:LYS:NZ	7:B:523:HOH:O	2.37	0.57
1:C:140:ILE:HD13	1:C:231:LYS:HB3	1.89	0.54
1:B:166:PHE:CE2	1:B:170:LYS:HD2	2.43	0.53
1:B:151:LYS:HE3	7:B:448:HOH:O	2.07	0.53
1:C:195:ASP:OD2	1:C:197[B]:LYS:HE3	2.08	0.53
1:B:136:VAL:O	1:B:147:SER:HA	2.10	0.51
1:C:60:LEU:HD13	1:C:99:GLN:HG3	1.93	0.51
1:A:38:HIS:CE1	1:A:206:LYS:H	2.30	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:133:ASP:OD1	1:C:151:LYS:HE2	2.13	0.49
1:C:182:VAL:O	1:C:185[B]:ARG:HD3	2.14	0.47
1:B:65:VAL:HG12	1:B:93:GLN:HB2	1.98	0.46
1:B:145:GLU:HB3	7:B:437:HOH:O	2.15	0.46
1:B:142:GLU:OE2	1:B:142:GLU:C	2.55	0.45
1:D:59:GLY:HA3	1:D:73:TRP:CZ2	2.52	0.45
1:B:128:PHE:CZ	1:B:219:THR:HA	2.51	0.45
1:D:84:LEU:HD22	1:D:84:LEU:N	2.32	0.45
1:A:60:LEU:HD11	1:A:120:LEU:HD13	1.99	0.44
1:B:84:LEU:N	1:B:84:LEU:HD12	2.33	0.44
1:A:177:LYS:HG3	7:A:508:HOH:O	2.18	0.44
1:A:55:LYS:HE2	7:A:475:HOH:O	2.18	0.44
1:A:66:THR:HG21	1:B:45:TYR:CE2	2.52	0.44
1:A:98:LYS:O	1:A:122[A]:LYS:HD3	2.18	0.44
1:C:39:ASN:N	7:C:510:HOH:O	2.52	0.43
1:D:83:LEU:C	1:D:84:LEU:HD22	2.39	0.43
1:B:59:GLY:HA3	1:B:73:TRP:CZ2	2.53	0.43
1:D:208:GLN:HG3	7:D:509:HOH:O	2.17	0.43
1:A:66:THR:HG21	1:B:45:TYR:CZ	2.52	0.43
1:D:60:LEU:HD13	1:D:99:GLN:HG3	2.01	0.42
1:A:208:GLN:HE22	2:E:1:NAG:H2	1.85	0.42
1:B:202:GLU:OE2	2:F:4:FUC:O3	2.33	0.42
1:C:168:LEU:HD23	1:C:168:LEU:HA	1.95	0.41
1:D:60:LEU:HG	1:D:72:VAL:HG12	2.01	0.41
1:A:38:HIS:HE1	1:A:206:LYS:H	1.69	0.40
1:A:128:PHE:CE1	1:A:219:THR:HG22	2.56	0.40
1:A:186:LYS:HE2	7:A:418[A]:HOH:O	2.20	0.40
1:A:208:GLN:NE2	1:A:210:HIS:HE1	2.20	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	199/218 (91%)	192 (96%)	7 (4%)	0	100	100
1	B	192/218 (88%)	187 (97%)	4 (2%)	1 (0%)	29	17
1	C	191/218 (88%)	184 (96%)	7 (4%)	0	100	100
1	D	195/218 (89%)	190 (97%)	5 (3%)	0	100	100
All	All	777/872 (89%)	753 (97%)	23 (3%)	1 (0%)	51	43

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	140	ILE

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	187/200 (94%)	187 (100%)	0	100	100
1	B	182/200 (91%)	179 (98%)	3 (2%)	62	52
1	C	181/200 (90%)	180 (99%)	1 (1%)	86	85
1	D	183/200 (92%)	181 (99%)	2 (1%)	73	67
All	All	733/800 (92%)	727 (99%)	6 (1%)	81	78

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	38	HIS
1	B	68	LYS
1	B	142	GLU
1	C	77	LYS
1	D	185	ARG
1	D	231	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	38	HIS
1	B	208	GLN
1	D	208	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

16 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	NAG	E	1	2	15,15,15	0.61	0	21,21,21	1.22	1 (4%)
2	GAL	E	2	2	11,11,12	0.54	0	15,15,17	1.15	2 (13%)
2	SIA	E	3	2	20,20,21	0.88	1 (5%)	24,28,31	1.03	2 (8%)
2	FUC	E	4	2	10,10,11	0.53	0	14,14,16	1.11	2 (14%)
2	NAG	F	1	2	15,15,15	0.59	0	21,21,21	1.01	1 (4%)
2	GAL	F	2	2	11,11,12	0.72	0	15,15,17	1.02	1 (6%)
2	SIA	F	3	2	20,20,21	1.13	2 (10%)	24,28,31	1.51	4 (16%)
2	FUC	F	4	2	10,10,11	0.77	0	14,14,16	1.04	1 (7%)
2	NAG	G	1	2	15,15,15	0.60	0	21,21,21	1.27	2 (9%)
2	GAL	G	2	2	11,11,12	0.57	0	15,15,17	1.02	0
2	SIA	G	3	2	20,20,21	0.96	2 (10%)	24,28,31	1.11	3 (12%)
2	FUC	G	4	2	10,10,11	0.78	0	14,14,16	1.01	1 (7%)
2	NAG	H	1	2	15,15,15	0.63	0	21,21,21	1.40	4 (19%)
2	GAL	H	2	2	11,11,12	0.63	0	15,15,17	0.73	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
2	SIA	H	3	2	20,20,21	0.94	1 (5%)	24,28,31	1.23	2 (8%)
2	FUC	H	4	2	10,10,11	0.94	1 (10%)	14,14,16	1.10	1 (7%)

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	NAG	E	1	2	-	0/6/26/26	0/1/1/1
2	GAL	E	2	2	-	0/2/19/22	0/1/1/1
2	SIA	E	3	2	-	0/18/34/38	0/1/1/1
2	FUC	E	4	2	-	-	0/1/1/1
2	NAG	F	1	2	-	0/6/26/26	0/1/1/1
2	GAL	F	2	2	-	0/2/19/22	0/1/1/1
2	SIA	F	3	2	-	0/18/34/38	0/1/1/1
2	FUC	F	4	2	-	-	0/1/1/1
2	NAG	G	1	2	-	0/6/26/26	0/1/1/1
2	GAL	G	2	2	-	0/2/19/22	0/1/1/1
2	SIA	G	3	2	-	0/18/34/38	0/1/1/1
2	FUC	G	4	2	-	-	0/1/1/1
2	NAG	H	1	2	-	0/6/26/26	0/1/1/1
2	GAL	H	2	2	-	0/2/19/22	0/1/1/1
2	SIA	H	3	2	-	0/18/34/38	0/1/1/1
2	FUC	H	4	2	-	-	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	3	SIA	O6-C2	-2.86	1.39	1.43
2	H	4	FUC	O5-C1	-2.59	1.39	1.43
2	E	3	SIA	O6-C2	-2.49	1.40	1.43
2	H	3	SIA	O6-C2	-2.43	1.40	1.43
2	G	3	SIA	O6-C2	-2.12	1.40	1.43
2	F	3	SIA	O6-C6	-2.08	1.40	1.44
2	G	3	SIA	C3-C2	2.07	1.55	1.52

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	1	NAG	C1-C2-N2	4.49	115.93	110.73
2	G	1	NAG	C1-C2-N2	-3.98	106.12	110.73

Continued on next page...

Continued from previous page...

| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	3	SIA	O6-C2-C3	-3.90	105.09	110.46
2	H	1	NAG	O5-C1-C2	3.58	113.12	109.52
2	H	3	SIA	O6-C2-C1	3.33	114.24	107.70
2	F	3	SIA	C6-O6-C2	3.01	117.78	111.34
2	H	4	FUC	O5-C1-C2	-2.75	106.53	110.77
2	E	2	GAL	O5-C5-C6	2.57	111.24	107.20
2	G	3	SIA	O1B-C1-C2	2.51	120.19	113.03
2	G	3	SIA	O1A-C1-C2	-2.49	116.69	122.57
2	H	1	NAG	C1-C2-N2	-2.44	107.90	110.73
2	E	3	SIA	O1A-C1-C2	-2.35	117.02	122.57
2	F	3	SIA	C6-C5-N5	-2.31	107.07	110.91
2	F	3	SIA	O6-C2-C1	2.29	112.19	107.70
2	F	2	GAL	C1-O5-C5	2.23	115.22	112.19
2	H	1	NAG	C6-C5-C4	2.20	118.16	113.00
2	H	3	SIA	O1B-C1-C2	2.17	119.21	113.03
2	H	1	NAG	C4-C3-C2	-2.15	107.19	110.34
2	E	4	FUC	C2-C3-C4	-2.15	107.18	110.89
2	G	3	SIA	C4-C3-C2	2.11	113.59	109.81
2	E	4	FUC	O2-C2-C1	2.11	113.47	109.15
2	G	4	FUC	C2-C3-C4	-2.09	107.28	110.89
2	F	4	FUC	O2-C2-C3	-2.08	105.96	110.14
2	G	1	NAG	O5-C5-C4	-2.07	105.93	109.69
2	F	1	NAG	O5-C1-C2	-2.07	107.44	109.52
2	E	3	SIA	O1B-C1-C2	2.02	118.79	113.03
2	E	2	GAL	O5-C1-C2	2.01	113.87	110.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	4	FUC	1	0
2	E	1	NAG	1	0
2	F	4	FUC	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.

Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain F

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain G

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain H

Bond lengths   Bond angles

Torsions   Rings
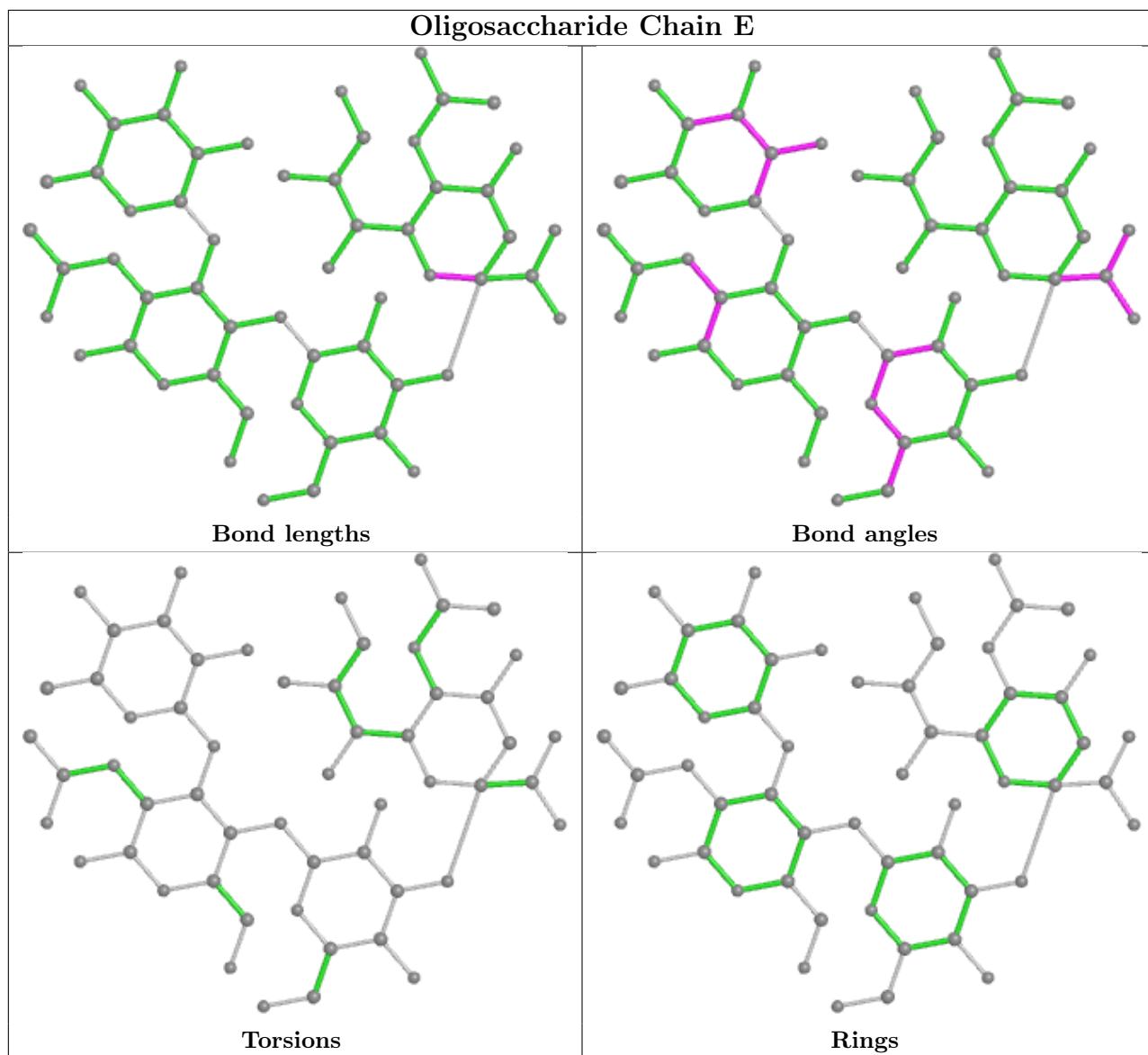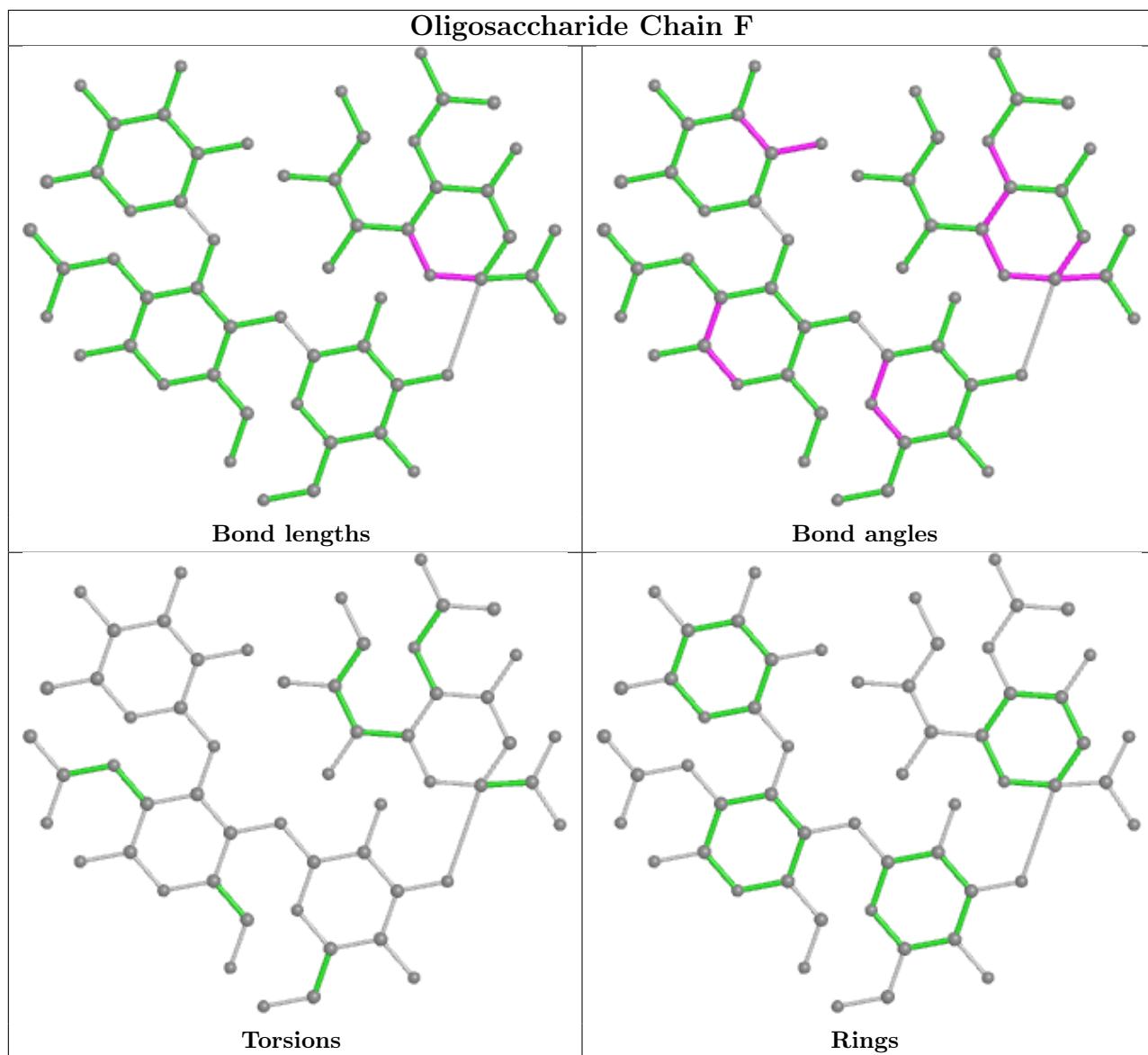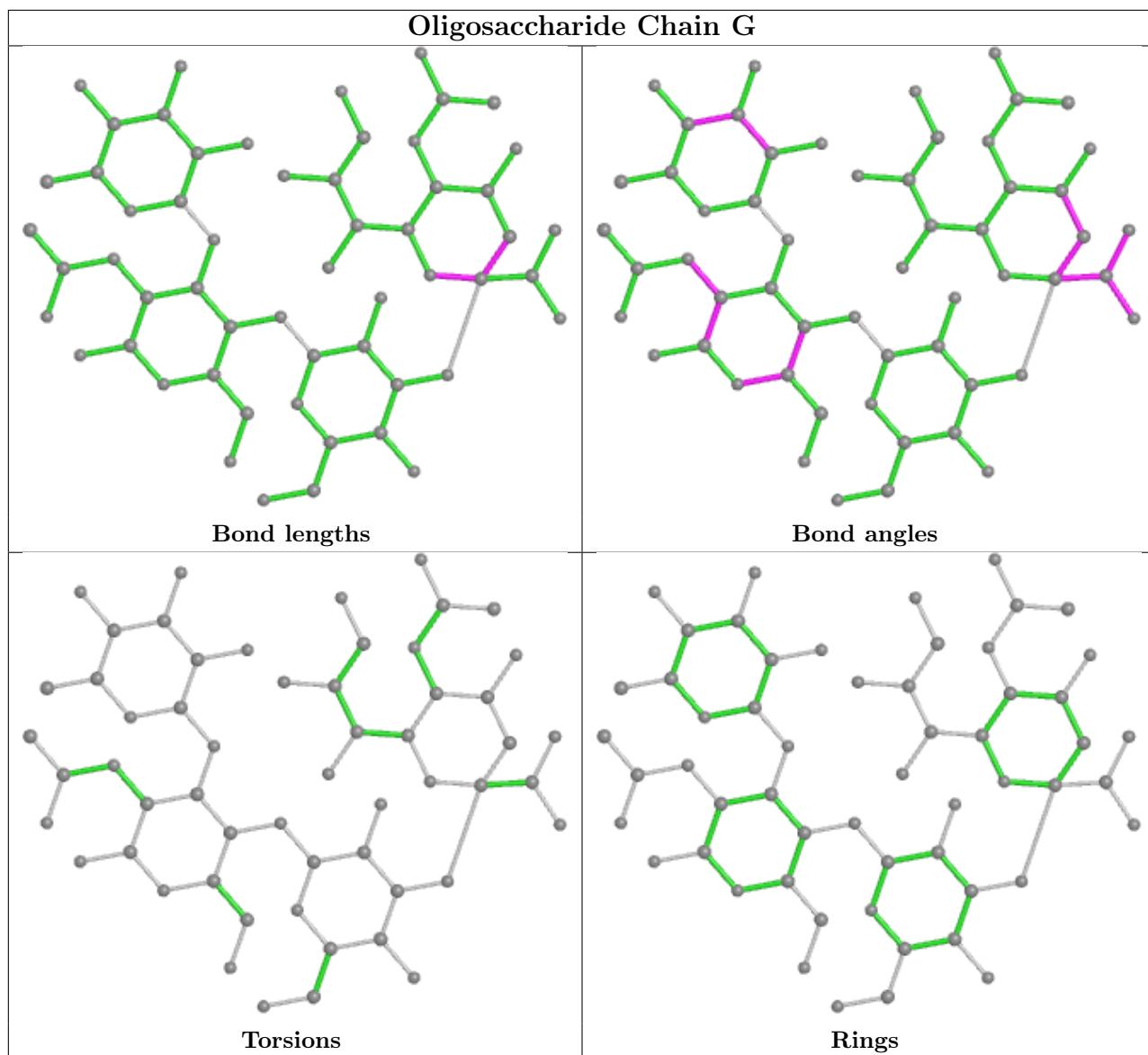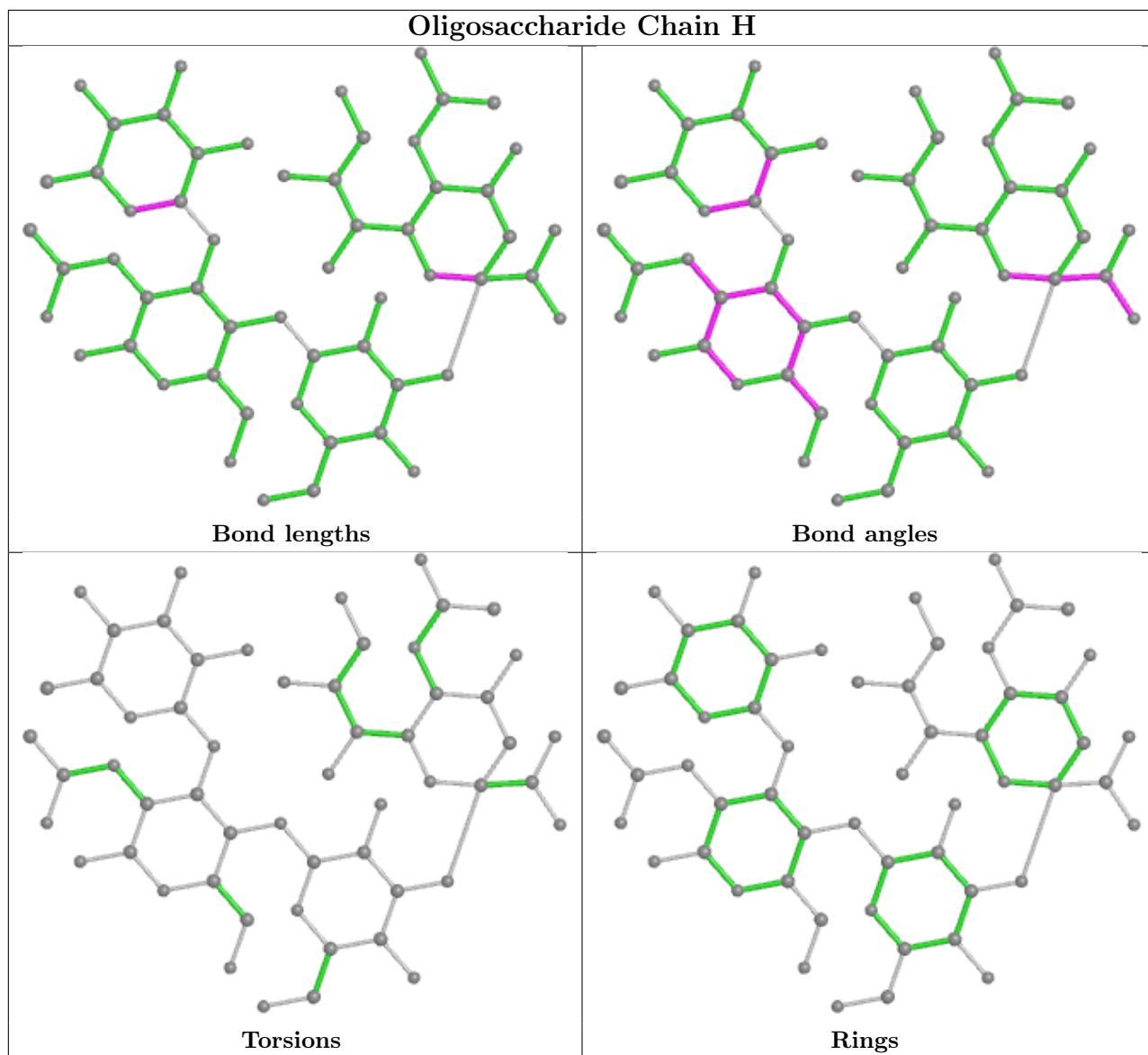
5.6 Ligand geometry ⓘ

Of 11 ligands modelled in this entry, 6 are monoatomic - leaving 5 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	#$|Z| > 2$
4	PEG	A	307	-	6,6,6	0.44	0	5,5,5	0.46	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
6	EDO	D	306	-	3,3,3	0.41	0	2,2,2	0.35	0
4	PEG	C	307	-	6,6,6	0.31	0	5,5,5	0.31	0
5	PGE	B	307	-	9,9,9	0.39	0	8,8,8	0.28	0
4	PEG	C	306	-	6,6,6	0.43	0	5,5,5	0.23	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
4	PEG	A	307	-	-	2/4/4/4	-
6	EDO	D	306	-	-	0/1/1/1	-
4	PEG	C	307	-	-	1/4/4/4	-
5	PGE	B	307	-	-	3/7/7/7	-
4	PEG	C	306	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	307	PGE	O2-C3-C4-O3
4	C	306	PEG	O2-C3-C4-O4
4	A	307	PEG	O2-C3-C4-O4
4	C	306	PEG	O1-C1-C2-O2
4	C	307	PEG	O2-C3-C4-O4
5	B	307	PGE	O3-C5-C6-O4
5	B	307	PGE	O1-C1-C2-O2
4	C	306	PEG	C4-C3-O2-C2
4	A	307	PEG	C4-C3-O2-C2

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

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	A	194/218 (88%)	0.04	5 (2%)	56	63	18, 30, 58, 79	0
1	B	192/218 (88%)	-0.03	8 (4%)	36	43	18, 30, 59, 83	0
1	C	191/218 (87%)	0.08	10 (5%)	27	34	18, 30, 61, 109	0
1	D	193/218 (88%)	0.15	10 (5%)	27	34	19, 33, 63, 80	0
All	All	770/872 (88%)	0.06	33 (4%)	35	42	18, 31, 61, 109	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	38	HIS	7.7
1	C	231	LYS	5.9
1	D	182	VAL	5.1
1	D	143	SER	5.1
1	C	147	SER	4.9
1	A	140	ILE	4.6
1	C	141	ASP	4.3
1	C	146	LYS	4.2
1	B	38	HIS	3.9
1	D	140	ILE	3.7
1	A	144	SER	3.3
1	C	140	ILE	3.3
1	C	142	GLU	3.2
1	B	140	ILE	3.1
1	B	231	LYS	3.0
1	C	149	ASP	3.0
1	A	143	SER	2.9
1	D	144	SER	2.7
1	A	141	ASP	2.6
1	D	76	ARG	2.6
1	D	146	LYS	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	65	VAL	2.5
1	C	139	LYS	2.4
1	B	142	GLU	2.4
1	C	76	ARG	2.3
1	B	63	GLN	2.3
1	B	146	LYS	2.2
1	D	141	ASP	2.2
1	D	145	GLU	2.2
1	B	67	ASP	2.1
1	D	183	ASN	2.1
1	C	129	SER	2.1
1	D	149	ASP	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

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	FUC	G	4	10/11	0.92	0.12	27,32,35,35	0
2	NAG	E	1	15/15	0.93	0.12	26,32,52,52	0
2	NAG	G	1	15/15	0.94	0.09	28,35,46,47	0
2	NAG	H	1	15/15	0.94	0.09	25,33,50,54	0
2	FUC	F	4	10/11	0.95	0.11	31,38,40,45	0
2	NAG	F	1	15/15	0.95	0.10	25,36,49,50	0
2	GAL	G	2	11/12	0.96	0.07	24,26,28,29	0
2	FUC	H	4	10/11	0.96	0.07	31,34,35,35	0
2	GAL	F	2	11/12	0.97	0.07	22,23,25,26	0
2	SIA	G	3	20/21	0.97	0.06	24,25,30,37	0
2	SIA	F	3	20/21	0.97	0.07	21,23,28,31	0
2	FUC	E	4	10/11	0.97	0.08	33,38,43,44	0
2	GAL	H	2	11/12	0.97	0.07	21,23,25,26	0
2	SIA	H	3	20/21	0.97	0.07	19,21,27,29	0
2	SIA	E	3	20/21	0.97	0.08	18,20,26,27	0
2	GAL	E	2	11/12	0.98	0.09	19,22,23,24	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:

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

Electron density around Chain F:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
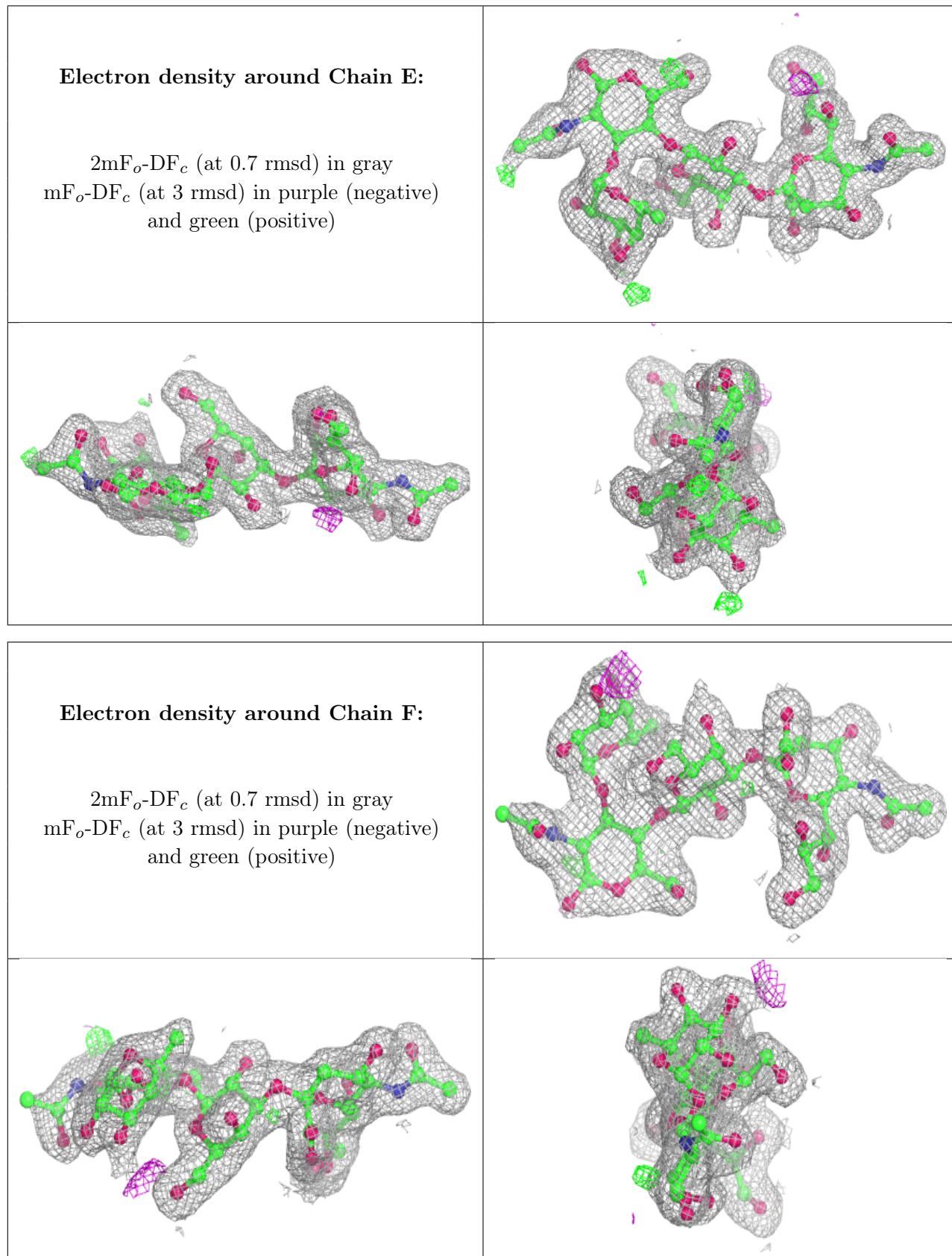
Electron density around Chain G:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
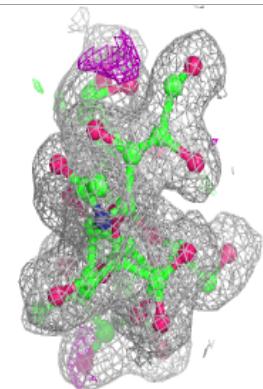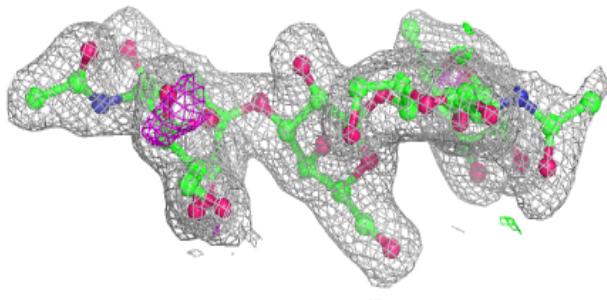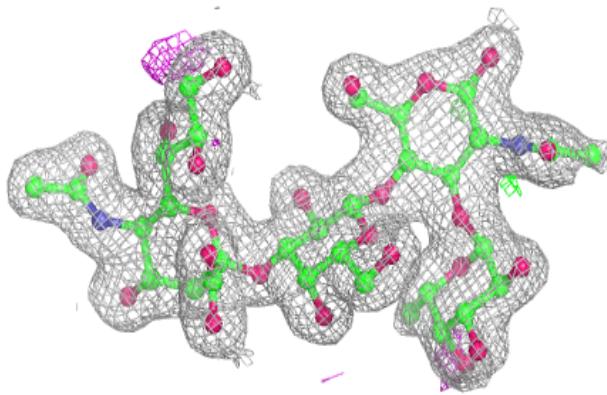
Electron density around Chain H:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
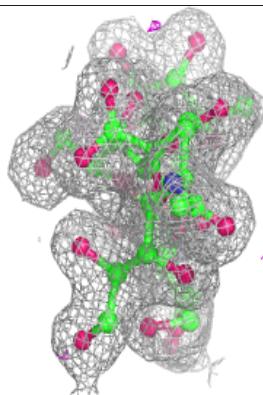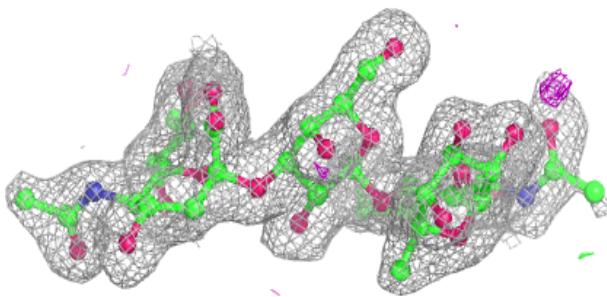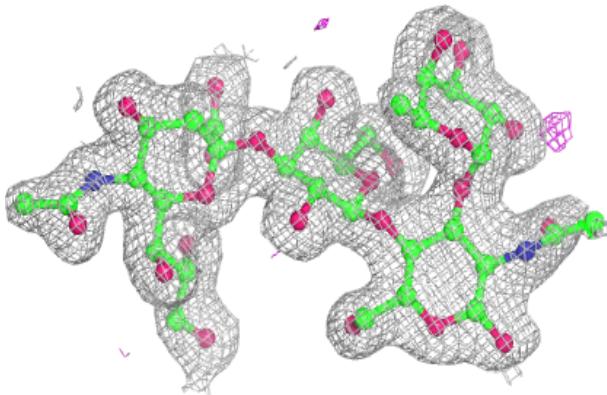
6.4 Ligands ⓘ

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
4	PEG	C	306	7/7	0.77	0.15	50,53,56,56	0
5	PGE	B	307	10/10	0.83	0.16	49,57,63,64	0
4	PEG	C	307	7/7	0.86	0.18	44,52,57,59	0
4	PEG	A	307	7/7	0.88	0.19	51,56,66,68	0
6	EDO	D	306	4/4	0.89	0.15	50,51,55,59	0
3	CL	A	305	1/1	0.95	0.07	41,41,41,41	0
3	CL	C	305	1/1	0.97	0.04	37,37,37,37	0
3	CL	B	306	1/1	0.97	0.05	33,33,33,33	0
3	CL	D	305	1/1	0.98	0.04	34,34,34,34	0
3	CL	A	306	1/1	0.99	0.05	31,31,31,31	0
3	CL	B	305	1/1	0.99	0.08	30,30,30,30	0

6.5 Other polymers ⓘ

There are no such residues in this entry.