Full wwPDB X-ray Structure Validation Report (i)

Oct 31, 2023 – 06:55 PM JST

PDB ID	:	5GX6
Title	:	Crystal structure of solute-binding protein complexed with unsaturated chondroitin disaccharide with a sulfate group at C-4 position of GalNAc
Authors	:	Oiki, S.; Mikami, B.; Murata, K.; Hashimoto, W.
Deposited on	:	2016-09-15
Resolution	:	1.81 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

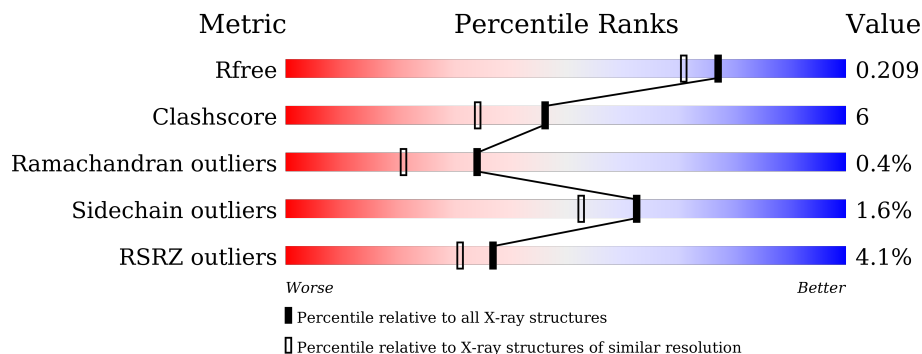|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 1.81 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		0.209
Clashscore		6
Ramachandran outliers		0.4%
Sidechain outliers		1.6%
RSRZ outliers		4.1%

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	7484 (1.84-1.80)
Clashscore	141614	8401 (1.84-1.80)
Ramachandran outliers	138981	8290 (1.84-1.80)
Sidechain outliers	138945	8290 (1.84-1.80)
RSRZ outliers	127900	7371 (1.84-1.80)

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	483	.% / 92% · · · 6% ·
1	B	483	5% / 89% · · · 8% · ·
1	C	483	6% / 89% · · · 7% · ·
2	D	2	50% · · · 50%
2	E	2	50% · · · 50%
2	F	2	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
4	EDO	B	608	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 12784 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.
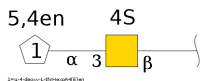
- Molecule 1 is a protein called Extracellular solute-binding protein family 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	474	Total	C	N	O	S	0	4	0
			3869	2485	641	737	6			
1	B	474	Total	C	N	O	S	0	11	0
			3928	2528	648	746	6			
1	C	474	Total	C	N	O	S	0	5	0
			3870	2488	638	737	7			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	18	MET	-	expression tag	UNP D1AWE0
B	18	MET	-	expression tag	UNP D1AWE0
C	18	MET	-	expression tag	UNP D1AWE0

- Molecule 2 is an oligosaccharide called 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1 -3)-2-acetamido-2-deoxy-4-O-sulfo-beta-D-galactopyranose.
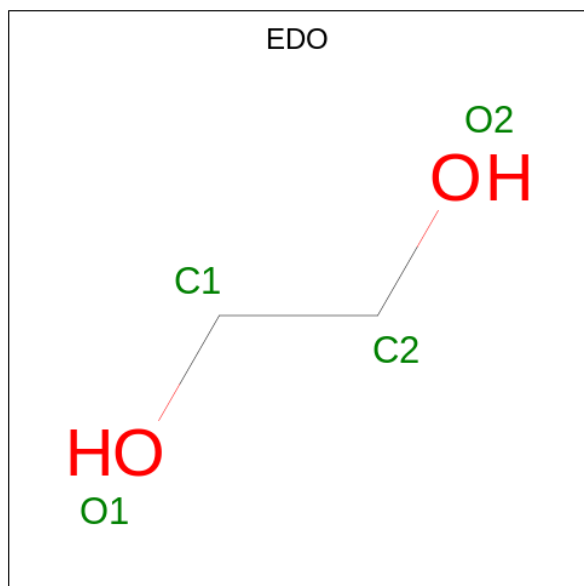

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	2	Total	C	N	O	S	0	0	0
			30	14	1	14	1			
2	E	2	Total	C	N	O	S	0	0	0
			30	14	1	14	1			
2	F	2	Total	C	N	O	S	0	0	0
			30	14	1	14	1			

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

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

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

EDO

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

Continued on next page...

Continued from previous page...

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

• Molecule 5 is water.

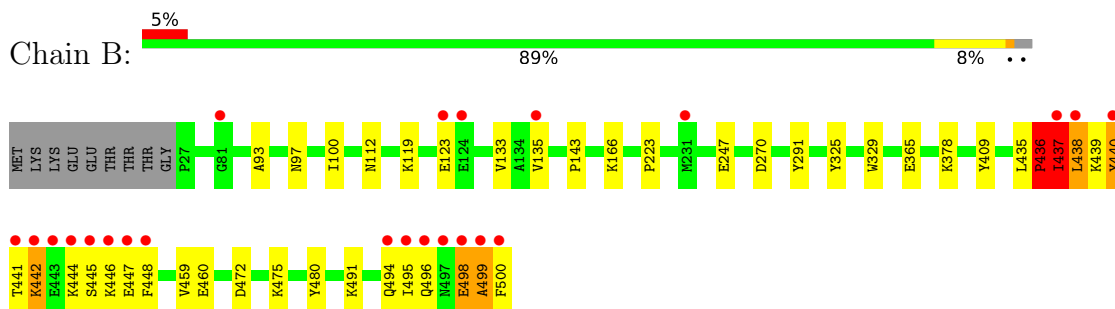| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	336	Total 336	O 336	0	0
5	B	346	Total 346	O 346	0	0
5	C	282	Total 282	O 282	0	0

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: Extracellular solute-binding protein family 1
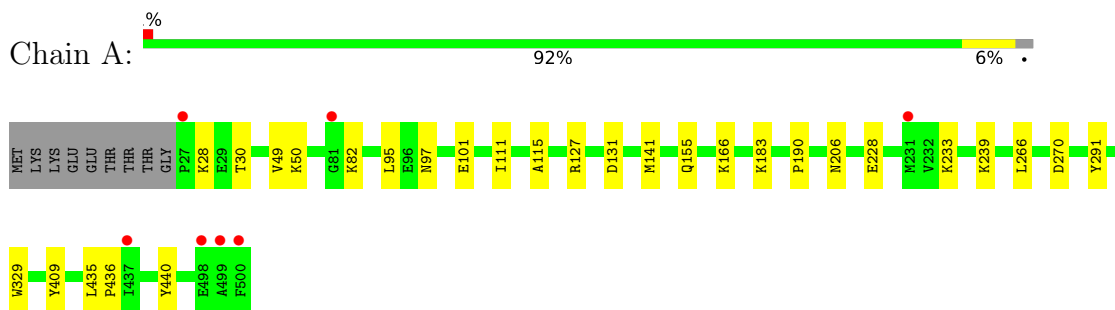
Chain A:

• Molecule 1: Extracellular solute-binding protein family 1

Chain B:

• Molecule 1: Extracellular solute-binding protein family 1

Chain C:

• Molecule 2: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-3)-2-acetamido-2-deoxy-4-O-sulfo-beta-D-galactopyranose

Chain D:

ASG1 GCD2
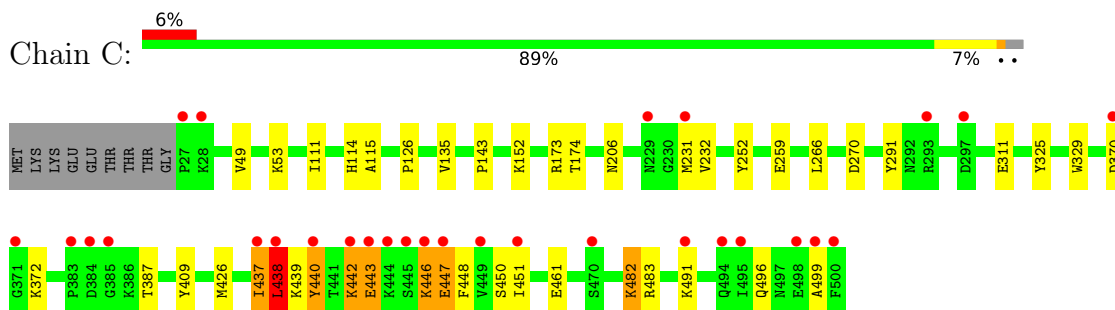
• Molecule 2: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-3)-2-acetamido-2-deoxy-4-O-sulfo-beta-D-galactopyranose

Chain E:

50%　　　　　　50%

ASG1 GCD2

• Molecule 2: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-3)-2-acetamido-2-deoxy-4-O-sulfo-beta-D-galactopyranose

Chain F:

100%

ASG1 GCD2

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	79.98Å 112.59Å 165.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.45 – 1.81 45.49 – 1.81	Depositor EDS
% Data completeness (in resolution range)	99.6 (39.45-1.81) 99.6 (45.49-1.81)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.73 (at 1.81Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.191 , 0.211 0.191 , 0.209	Depositor DCC
R_{free} test set	6848 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	17.5	Xtriage
Anisotropy	0.354	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 45.2	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12784	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.12% 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: ASG, CA, GCD, 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	A	0.39	1/3959 (0.0%)	0.54	1/5356 (0.0%)
1	B	0.50	0/4019	0.66	3/5435 (0.1%)
1	C	0.42	0/3963	0.59	2/5360 (0.0%)
All	All	0.44	1/11941 (0.0%)	0.60	6/16151 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	436	PRO	N-CD	5.11	1.55	1.47

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	436	PRO	O-C-N	-13.49	101.11	122.70
1	B	436	PRO	CA-C-N	9.44	137.97	117.20
1	C	491	LYS	CD-CE-NZ	-9.15	90.65	111.70
1	B	435	LEU	C-N-CD	6.07	141.15	128.40
1	C	438	LEU	CB-CA-C	5.76	121.14	110.20
1	A	435	LEU	C-N-CD	5.04	138.98	128.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	436	PRO	Mainchain,Peptide
1	B	446[A]	LYS	Peptide

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	3869	0	3813	16	0
1	B	3928	0	3894	101	0
1	C	3870	0	3821	32	0
2	D	30	0	14	0	0
2	E	30	0	14	0	0
2	F	30	0	14	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	20	0	30	2	0
4	B	28	0	42	10	0
4	C	12	0	18	2	0
5	A	336	0	0	2	1
5	B	346	0	0	2	0
5	C	282	0	0	4	1
All	All	12784	0	11660	148	1

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 6.

All (148) 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:B:438[B]:LEU:HD23	1:B:499:ALA:CB	1.40	1.51
1:B:438[B]:LEU:CD2	1:B:500:PHE:H	1.34	1.41
1:B:438[B]:LEU:HD21	1:B:500:PHE:N	1.42	1.34
1:B:438[B]:LEU:CD2	1:B:499:ALA:HB1	1.61	1.30
1:B:447:GLU:OE2	1:B:491:LYS:NZ	1.65	1.29
1:B:444[B]:LYS:NZ	1:B:498:GLU:HG2	1.45	1.28
1:B:438[A]:LEU:HB3	1:B:500:PHE:N	1.49	1.26

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:437[B]:ILE:HG21	5:B:846:HOH:O	1.26	1.26
1:B:438[B]:LEU:CD2	1:B:499:ALA:CB	2.11	1.25
1:B:438[A]:LEU:C	1:B:499:ALA:HB1	1.66	1.15
1:B:438[A]:LEU:HD13	1:B:500:PHE:HB2	1.28	1.11
1:B:438[A]:LEU:CD2	1:B:500:PHE:HA	1.82	1.09
1:B:438[A]:LEU:CB	1:B:500:PHE:H	1.58	1.07
1:B:438[A]:LEU:HD22	1:B:500:PHE:CA	1.84	1.07
1:B:444[B]:LYS:HZ3	1:B:498:GLU:HG2	0.97	1.00
1:B:440[B]:TYR:HB2	1:B:445[B]:SER:HB2	1.44	0.99
1:B:438[A]:LEU:HD22	1:B:500:PHE:HA	0.98	0.97
1:B:438[A]:LEU:CB	1:B:499:ALA:CB	2.42	0.96
1:B:438[B]:LEU:CD2	1:B:499:ALA:HB3	1.92	0.96
1:B:440[A]:TYR:HD2	1:B:498:GLU:O	1.48	0.94
1:B:444[B]:LYS:NZ	1:B:498:GLU:CG	2.30	0.94
1:B:438[A]:LEU:HB3	1:B:499:ALA:CB	2.00	0.92
1:B:438[B]:LEU:HD22	1:B:499:ALA:HB3	1.51	0.92
1:B:97:ASN:OD1	1:B:445[B]:SER:OG	1.87	0.90
1:B:438[A]:LEU:CB	1:B:499:ALA:HB1	2.00	0.90
1:B:438[A]:LEU:C	1:B:499:ALA:CB	2.41	0.88
1:B:440[A]:TYR:HD1	1:B:441[A]:THR:HG23	1.39	0.86
1:B:444[A]:LYS:HB3	1:B:495:ILE:HG23	1.57	0.85
1:B:444[A]:LYS:HD3	1:B:498:GLU:HG2	1.57	0.84
1:B:438[A]:LEU:HB2	1:B:499:ALA:HB3	1.57	0.83
1:B:438[A]:LEU:CA	1:B:499:ALA:HB1	2.08	0.83
1:B:438[B]:LEU:CD2	1:B:500:PHE:N	2.16	0.82
1:B:447:GLU:HG2	1:B:495:ILE:HD11	1.61	0.81
1:B:438[A]:LEU:CB	1:B:499:ALA:HB3	2.07	0.81
1:B:438[A]:LEU:HD13	1:B:500:PHE:CB	2.11	0.80
1:C:174:THR:HG22	1:C:311[C]:GLU:HG2	1.63	0.79
1:B:440[A]:TYR:CD2	1:B:498:GLU:O	2.35	0.78
1:B:444[B]:LYS:HZ3	1:B:498:GLU:CG	1.90	0.78
1:B:436:PRO:O	1:B:438[B]:LEU:HD12	1.84	0.77
1:B:439[A]:LYS:HE2	4:B:608:EDO:H11	1.64	0.77
1:B:440[A]:TYR:CD1	1:B:441[A]:THR:HG23	2.20	0.74
1:B:438[A]:LEU:HB3	1:B:499:ALA:HB1	1.65	0.73
1:B:448:PHE:HB2	1:B:495:ILE:HG21	1.69	0.73
1:B:444[B]:LYS:O	1:B:447:GLU:N	2.21	0.73
1:B:438[B]:LEU:HD21	1:B:500:PHE:H	0.60	0.72
1:B:438[A]:LEU:CD1	1:B:500:PHE:HB2	2.14	0.72
1:B:440[A]:TYR:HD1	1:B:441[A]:THR:CG2	2.02	0.72
1:C:447:GLU:O	1:C:450:SER:HB3	1.90	0.72

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:447:GLU:CD	1:B:491:LYS:NZ	2.43	0.71
1:A:270[A]:ASP:OD1	1:A:291:TYR:OH	2.06	0.71
1:B:440[B]:TYR:CB	1:B:445[B]:SER:HB2	2.20	0.69
1:B:447:GLU:HG2	1:B:495:ILE:CD1	2.21	0.69
1:C:461:GLU:OE1	1:C:483:ARG:NH2	2.26	0.69
1:B:93:ALA:HA	4:B:608:EDO:H21	1.74	0.68
1:B:444[B]:LYS:HZ1	1:B:498:GLU:HG2	1.54	0.66
1:B:439[B]:LYS:HD2	4:B:608:EDO:H11	1.78	0.65
1:B:437[B]:ILE:HG22	1:B:438[B]:LEU:N	2.13	0.64
1:C:135:VAL:HG11	1:C:440:TYR:O	1.97	0.64
1:B:438[A]:LEU:O	1:B:499:ALA:CB	2.45	0.64
1:C:438:LEU:CD1	1:C:499:ALA:O	2.47	0.63
1:B:438[B]:LEU:HD21	1:B:500:PHE:CA	2.27	0.62
1:B:439[A]:LYS:N	1:B:499:ALA:HB1	2.14	0.62
1:C:442:LYS:HG2	1:C:443:GLU:OE1	2.00	0.62
1:B:438[B]:LEU:HD23	1:B:499:ALA:HB1	0.66	0.60
1:C:270[A]:ASP:OD1	1:C:291:TYR:OH	2.16	0.59
1:C:259:GLU:HB3	4:C:604:EDO:H12	1.85	0.59
1:A:155:GLN:HG3	4:A:605:EDO:H21	1.85	0.58
1:C:438:LEU:HD11	1:C:499:ALA:O	2.03	0.58
1:B:448:PHE:HB2	1:B:495:ILE:CG2	2.33	0.58
1:B:448:PHE:CE1	1:B:496:GLN:HG2	2.40	0.56
1:B:447:GLU:CD	1:B:491:LYS:HZ2	2.09	0.55
1:B:438[A]:LEU:HB2	1:B:499:ALA:CB	2.23	0.54
1:B:133:VAL:O	1:B:439[B]:LYS:HE3	2.07	0.53
1:A:183:LYS:HE2	5:A:1259:HOH:O	2.09	0.53
1:B:123:GLU:O	1:C:126:PRO:HG3	2.08	0.53
1:B:447:GLU:CG	1:B:495:ILE:HD11	2.37	0.53
1:A:228:GLU:O	1:A:228:GLU:HG3	2.09	0.52
1:B:498:GLU:OE2	1:B:498:GLU:HA	2.09	0.52
1:B:438[A]:LEU:CB	1:B:500:PHE:N	2.38	0.52
1:B:437[A]:ILE:HG23	5:B:757:HOH:O	2.10	0.52
1:B:437[A]:ILE:HG13	1:B:438[A]:LEU:N	2.22	0.51
1:B:100:ILE:CG2	1:B:442[B]:LYS:HG2	2.41	0.51
1:B:444[A]:LYS:O	1:B:495:ILE:HG21	2.10	0.51
1:C:448:PHE:O	1:C:451:ILE:N	2.44	0.50
1:C:325:TYR:OH	1:C:437:ILE:HG12	2.11	0.50
1:B:135:VAL:HG21	1:B:440[B]:TYR:C	2.32	0.50
1:B:444[B]:LYS:O	1:B:448:PHE:N	2.44	0.49
1:C:152:LYS:HE3	1:C:426[B]:MET:HE1	1.95	0.49
1:A:206:ASN:HB3	1:A:266:LEU:HG	1.95	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:372:LYS:HA	1:C:372:LYS:HD3	1.63	0.48
1:C:370:ASP:HB2	5:C:945:HOH:O	2.13	0.48
1:C:482:LYS:N	1:C:482:LYS:HD3	2.29	0.48
1:B:444[A]:LYS:HD3	1:B:498:GLU:CG	2.37	0.48
1:A:228:GLU:HG2	1:A:233:LYS:HD3	1.96	0.47
1:B:93:ALA:CA	4:B:608:EDO:H21	2.43	0.47
1:B:440[A]:TYR:CD1	1:B:441[A]:THR:CG2	2.88	0.47
1:B:100:ILE:HG23	1:B:442[B]:LYS:HG2	1.97	0.47
1:C:49:VAL:HG12	1:C:53:LYS:HE2	1.97	0.47
1:B:494:GLN:O	1:B:498:GLU:N	2.42	0.47
1:B:438[A]:LEU:HD22	1:B:500:PHE:N	2.27	0.46
1:A:49:VAL:HG22	4:A:607:EDO:H22	1.97	0.46
1:A:97:ASN:O	1:A:101:GLU:HG3	2.15	0.46
1:C:114:HIS:HD2	5:C:957:HOH:O	1.98	0.46
1:B:437[B]:ILE:O	1:B:438[B]:LEU:HB2	2.16	0.46
1:C:174:THR:HG21	5:C:715:HOH:O	2.15	0.46
1:B:270:ASP:OD1	1:B:291:TYR:OH	2.25	0.46
1:A:95:LEU:HB3	1:A:141:MET:HE1	1.98	0.45
1:B:494:GLN:O	1:B:495:ILE:C	2.51	0.45
1:C:387:THR:HB	5:C:909:HOH:O	2.16	0.45
1:C:448:PHE:CZ	1:C:496:GLN:NE2	2.85	0.45
1:B:439[B]:LYS:CD	4:B:608:EDO:H11	2.40	0.45
1:B:472:ASP:OD2	1:B:475:LYS:HE3	2.17	0.45
1:C:173:ARG:O	1:C:173:ARG:HG3	2.17	0.45
1:B:438[A]:LEU:O	1:B:499:ALA:HB2	2.16	0.44
1:B:442[A]:LYS:HD2	1:B:442[A]:LYS:C	2.37	0.44
1:B:438[A]:LEU:CG	1:B:500:PHE:H	2.22	0.44
1:C:143:PRO:HB3	1:C:325:TYR:CD2	2.51	0.44
1:C:438:LEU:HD12	1:C:499:ALA:O	2.17	0.44
1:B:143:PRO:HB3	1:B:325:TYR:CD2	2.52	0.43
1:B:480:TYR:HB2	4:B:609:EDO:H12	2.00	0.43
1:C:259:GLU:H	4:C:604:EDO:C1	2.30	0.43
1:B:365:GLU:OE2	1:B:378:LYS:NZ	2.44	0.43
1:B:438[A]:LEU:CD2	1:B:500:PHE:CA	2.67	0.43
1:B:439[B]:LYS:C	1:B:440[B]:TYR:CG	2.91	0.43
1:C:152:LYS:HE3	1:C:426[B]:MET:CE	2.48	0.43
1:A:166:LYS:CD	1:A:190:PRO:O	2.67	0.43
1:B:112:ASN:HA	1:B:119:LYS:HE3	2.01	0.43
1:A:82:LYS:HE2	5:A:1062:HOH:O	2.18	0.43
1:C:231:MET:HG2	1:C:232:VAL:N	2.34	0.43
1:B:223:PRO:O	4:B:606:EDO:H11	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:446:LYS:HE2	1:C:446:LYS:HB2	1.75	0.42
1:B:439[B]:LYS:O	1:B:440[B]:TYR:CD2	2.73	0.42
1:C:111:ILE:HA	1:C:115:ALA:HB3	2.01	0.42
1:B:438[A]:LEU:HB3	1:B:499:ALA:C	2.28	0.42
1:B:440[A]:TYR:CD1	1:B:440[A]:TYR:C	2.93	0.42
1:C:252:TYR:CE1	1:C:259:GLU:HA	2.55	0.42
1:A:111:ILE:HA	1:A:115:ALA:HB3	2.02	0.41
1:A:127:ARG:NH2	1:A:131:ASP:OD1	2.50	0.41
1:A:239:LYS:HD3	1:A:239:LYS:HA	1.88	0.41
1:B:460:GLU:HG3	4:B:606:EDO:H22	2.03	0.41
1:B:442[A]:LYS:HD2	1:B:442[A]:LYS:O	2.21	0.41
1:C:442:LYS:HE2	1:C:442:LYS:HB2	1.76	0.41
1:A:166:LYS:NZ	1:A:190:PRO:O	2.37	0.41
1:B:459:VAL:HG11	4:B:606:EDO:H12	2.03	0.41
1:C:206:ASN:HB3	1:C:266:LEU:HG	2.03	0.41
1:A:28:LYS:HE3	1:A:30:THR:OG1	2.20	0.40
1:B:491:LYS:HD2	1:B:491:LYS:HA	1.76	0.40
1:B:247:GLU:OE2	4:B:605:EDO:O1	2.33	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:1001:HOH:O	5:C:715:HOH:O[3_654]	1.76	0.44

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	476/483 (99%)	469 (98%)	7 (2%)	0	100	100
1	B	483/483 (100%)	465 (96%)	11 (2%)	7 (1%)	11	3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	477/483 (99%)	465 (98%)	10 (2%)	2 (0%)	34	21
All	All	1436/1449 (99%)	1399 (97%)	28 (2%)	9 (1%)	34	12

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	437[A]	ILE
1	B	437[B]	ILE
1	B	438[A]	LEU
1	B	438[B]	LEU
1	B	499	ALA
1	B	440[A]	TYR
1	B	440[B]	TYR
1	C	442	LYS
1	C	447	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	412/416 (99%)	408 (99%)	4 (1%)	76	70
1	B	419/416 (101%)	411 (98%)	8 (2%)	57	45
1	C	413/416 (99%)	404 (98%)	9 (2%)	52	39
All	All	1244/1248 (100%)	1223 (98%)	21 (2%)	62	50

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	50	LYS
1	A	329	TRP
1	A	409	TYR
1	A	440	TYR
1	B	166	LYS
1	B	329	TRP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	409	TYR
1	B	437[A]	ILE
1	B	437[B]	ILE
1	B	442[A]	LYS
1	B	442[B]	LYS
1	B	498	GLU
1	C	329	TRP
1	C	409	TYR
1	C	437	ILE
1	C	438	LEU
1	C	439	LYS
1	C	440	TYR
1	C	443	GLU
1	C	446	LYS
1	C	482	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	114	HIS
1	C	161	GLN
1	C	464	GLN
1	C	486	ASN

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 ⓘ

6 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond

length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	ASG	D	1	2	19,19,19	0.98	0	23,28,28	0.88	0
2	GCD	D	2	2	10,11,12	3.30	3 (30%)	13,15,17	1.55	3 (23%)
2	ASG	E	1	2	19,19,19	1.06	0	23,28,28	0.87	0
2	GCD	E	2	2	10,11,12	3.32	3 (30%)	13,15,17	1.64	3 (23%)
2	ASG	F	1	2	19,19,19	1.14	1 (5%)	23,28,28	0.89	0
2	GCD	F	2	2	10,11,12	3.30	3 (30%)	13,15,17	1.46	3 (23%)

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	ASG	D	1	2	-	3/11/31/31	0/1/1/1
2	GCD	D	2	2	-	0/4/17/20	0/1/1/1
2	ASG	E	1	2	-	3/11/31/31	0/1/1/1
2	GCD	E	2	2	-	0/4/17/20	0/1/1/1
2	ASG	F	1	2	-	5/11/31/31	0/1/1/1
2	GCD	F	2	2	-	0/4/17/20	0/1/1/1

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	2	GCD	C3-C4	-8.46	1.39	1.50
2	E	2	GCD	C3-C4	-8.45	1.39	1.50
2	D	2	GCD	C3-C4	-8.27	1.40	1.50
2	E	2	GCD	C4-C5	4.58	1.40	1.33
2	D	2	GCD	C4-C5	4.56	1.40	1.33
2	F	2	GCD	C4-C5	4.25	1.39	1.33
2	D	2	GCD	C5-C6	-3.85	1.39	1.48
2	F	2	GCD	C5-C6	-3.80	1.39	1.48
2	E	2	GCD	C5-C6	-3.24	1.40	1.48
2	F	1	ASG	C1-C2	-3.11	1.49	1.52

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	2	GCD	O5-C5-C4	-2.87	122.39	124.81
2	D	2	GCD	O6B-C6-C5	2.76	121.09	114.20
2	F	2	GCD	O6B-C6-C5	2.68	120.89	114.20
2	D	2	GCD	C2-C3-C4	-2.50	108.90	112.32
2	D	2	GCD	O3-C3-C2	-2.31	105.41	109.42
2	E	2	GCD	O6B-C6-O6A	-2.24	118.49	123.61
2	E	2	GCD	O6B-C6-C5	2.13	119.52	114.20
2	F	2	GCD	C2-C3-C4	-2.09	109.46	112.32
2	F	2	GCD	O5-C5-C4	-2.03	123.10	124.81

There are no chirality outliers.

All (11) torsion outliers are listed below:

| Mol	Chain	Res	Type	Atoms
2	D	1	ASG	C4-O4-S-OSB
2	E	1	ASG	C4-O4-S-OSA
2	E	1	ASG	C4-O4-S-OSB
2	F	1	ASG	C4-O4-S-OSB
2	F	1	ASG	C8-C7-N2-C2
2	F	1	ASG	O7-C7-N2-C2
2	D	1	ASG	C4-O4-S-OSC
2	E	1	ASG	C4-O4-S-OSC
2	F	1	ASG	C4-O4-S-OSC
2	D	1	ASG	C4-O4-S-OSA
2	F	1	ASG	C4-O4-S-OSA

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.

Oligosaccharide Chain D

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain F
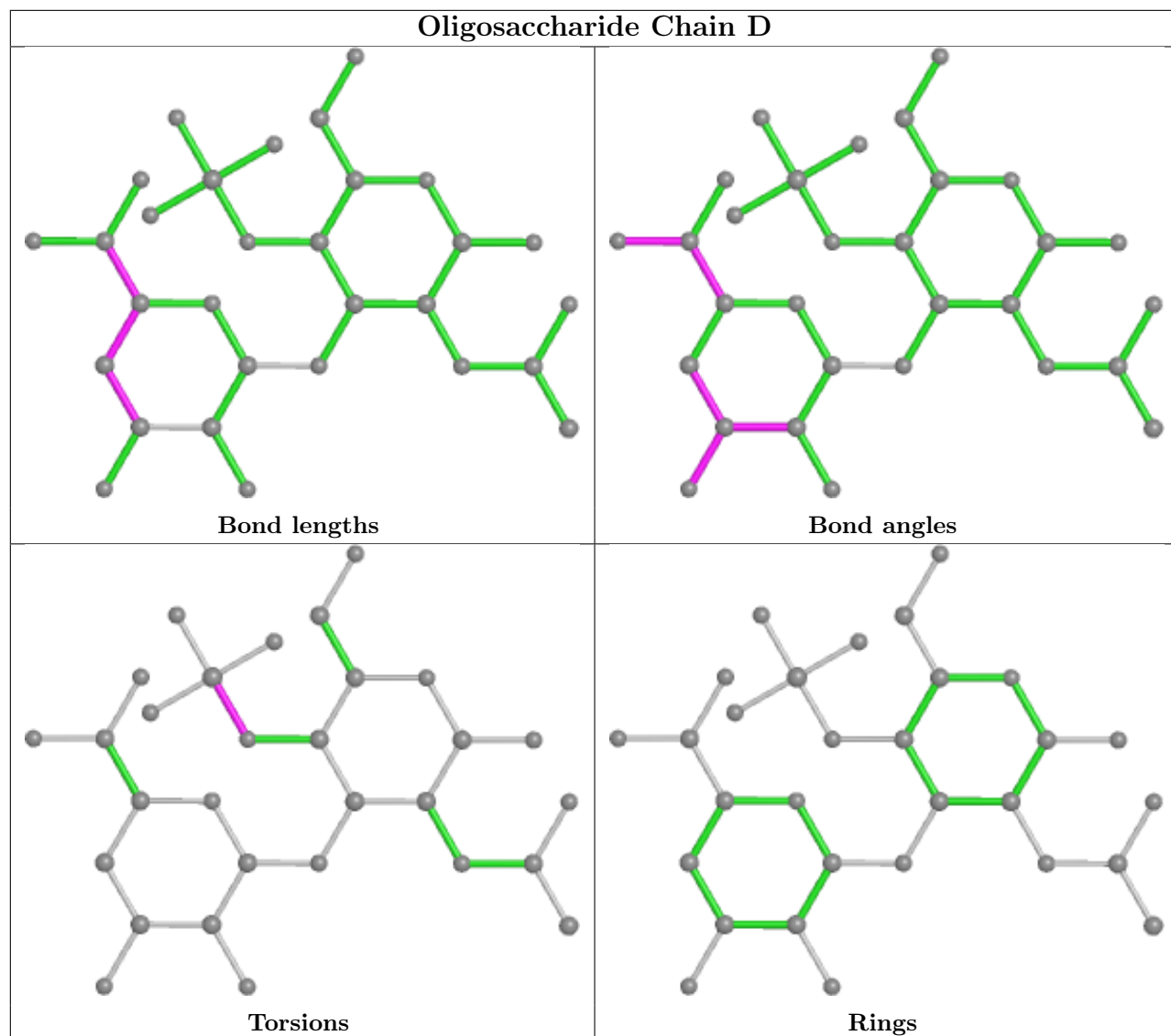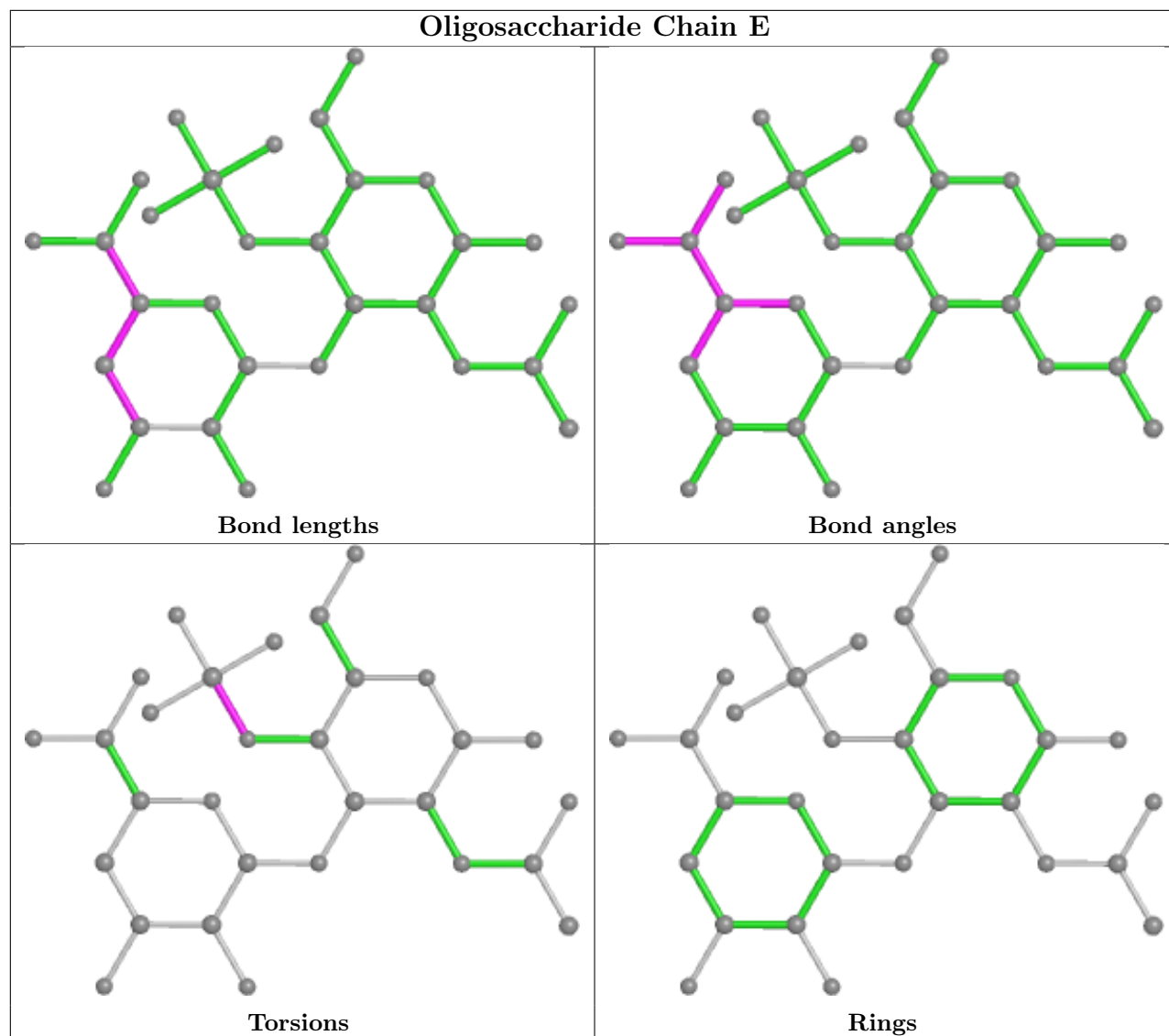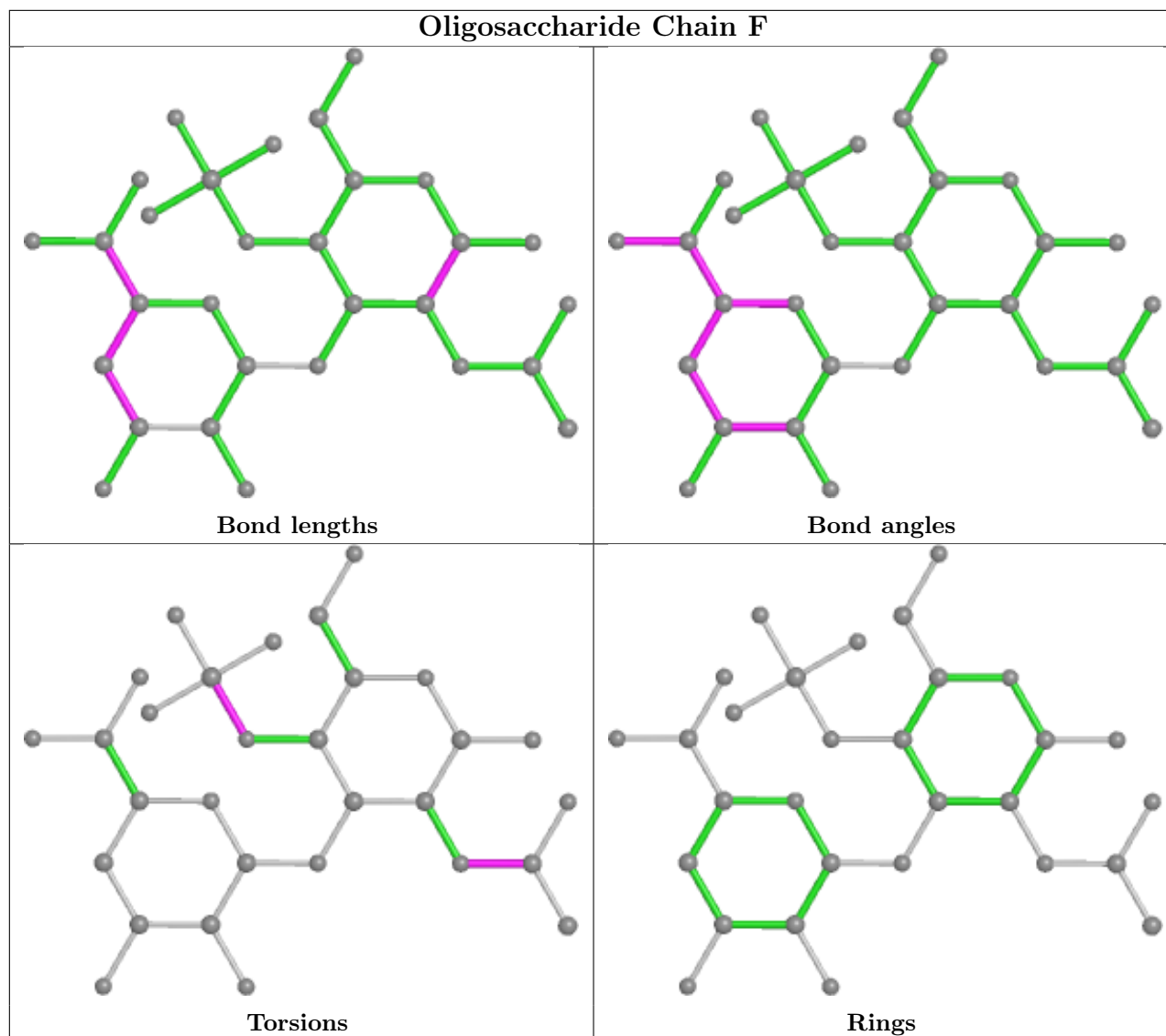
5.6 Ligand geometry ⓘ

Of 18 ligands modelled in this entry, 3 are monoatomic - leaving 15 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	EDO	B	605	-	3,3,3	0.44	0	2,2,2	0.34	0
4	EDO	B	607	-	3,3,3	0.46	0	2,2,2	0.26	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\lvert Z\rvert > 2$	Counts	RMSZ	$\#\lvert Z\rvert > 2$
4	EDO	C	604	-	3,3,3	0.36	0	2,2,2	0.49	0
4	EDO	C	605	-	3,3,3	0.45	0	2,2,2	0.31	0
4	EDO	A	604	-	3,3,3	0.44	0	2,2,2	0.41	0
4	EDO	B	606	-	3,3,3	0.35	0	2,2,2	0.16	0
4	EDO	B	609	-	3,3,3	0.48	0	2,2,2	0.13	0
4	EDO	C	606	-	3,3,3	0.51	0	2,2,2	0.17	0
4	EDO	A	608	-	3,3,3	0.43	0	2,2,2	0.42	0
4	EDO	A	607	-	3,3,3	0.42	0	2,2,2	0.36	0
4	EDO	B	608	-	3,3,3	0.40	0	2,2,2	0.26	0
4	EDO	A	606	-	3,3,3	0.43	0	2,2,2	0.20	0
4	EDO	B	610	-	3,3,3	0.42	0	2,2,2	0.51	0
4	EDO	B	604	-	3,3,3	0.51	0	2,2,2	0.42	0
4	EDO	A	605	-	3,3,3	0.37	0	2,2,2	0.41	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	EDO	B	605	-	-	0/1/1/1	-
4	EDO	B	607	-	-	1/1/1/1	-
4	EDO	C	604	-	-	0/1/1/1	-
4	EDO	C	605	-	-	0/1/1/1	-
4	EDO	A	604	-	-	0/1/1/1	-
4	EDO	B	606	-	-	0/1/1/1	-
4	EDO	B	609	-	-	1/1/1/1	-
4	EDO	C	606	-	-	1/1/1/1	-
4	EDO	A	608	-	-	1/1/1/1	-
4	EDO	A	607	-	-	1/1/1/1	-
4	EDO	B	608	-	-	1/1/1/1	-
4	EDO	A	606	-	-	0/1/1/1	-
4	EDO	B	610	-	-	1/1/1/1	-
4	EDO	B	604	-	-	0/1/1/1	-
4	EDO	A	605	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	607	EDO	O1-C1-C2-O2
4	A	608	EDO	O1-C1-C2-O2
4	B	609	EDO	O1-C1-C2-O2
4	C	606	EDO	O1-C1-C2-O2
4	B	608	EDO	O1-C1-C2-O2
4	B	607	EDO	O1-C1-C2-O2
4	B	610	EDO	O1-C1-C2-O2

There are no ring outliers.

7 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	605	EDO	1	0
4	C	604	EDO	2	0
4	B	606	EDO	3	0
4	B	609	EDO	1	0
4	A	607	EDO	1	0
4	B	608	EDO	5	0
4	A	605	EDO	1	0

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	A	474/483 (98%)	-0.08	7 (1%)	73	70	10, 17, 31, 43	0
1	B	474/483 (98%)	0.06	23 (4%)	29	24	10, 17, 37, 55	0
1	C	474/483 (98%)	0.16	29 (6%)	21	16	11, 19, 44, 64	0
All	All	1422/1449 (98%)	0.05	59 (4%)	37	31	10, 18, 37, 64	0

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	500	PHE	9.7
1	B	440[A]	TYR	7.7
1	B	499	ALA	7.7
1	B	498	GLU	6.0
1	C	500	PHE	6.0
1	C	385	GLY	5.6
1	C	440	TYR	5.4
1	C	437	ILE	4.5
1	B	445[A]	SER	4.4
1	B	446[A]	LYS	4.3
1	C	498	GLU	4.3
1	C	443	GLU	4.2
1	C	499	ALA	4.1
1	C	231	MET	4.1
1	C	449	VAL	4.1
1	B	438[A]	LEU	3.9
1	C	446	LYS	3.9
1	C	442	LYS	3.8
1	B	443[A]	GLU	3.8
1	B	496	GLN	3.8
1	B	495	ILE	3.7
1	B	494	GLN	3.7
1	B	437[A]	ILE	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	135	VAL	3.6
1	A	231	MET	3.4
1	A	81	GLY	3.4
1	C	384	ASP	3.4
1	C	445	SER	3.4
1	C	491	LYS	3.3
1	B	447	GLU	3.3
1	A	498	GLU	3.2
1	A	27	PRO	3.2
1	B	442[A]	LYS	3.2
1	C	27	PRO	3.1
1	C	438	LEU	3.0
1	C	370	ASP	2.9
1	C	494	GLN	2.8
1	B	123	GLU	2.7
1	A	499	ALA	2.6
1	B	497	ASN	2.6
1	C	28	LYS	2.6
1	C	444	LYS	2.6
1	B	231	MET	2.5
1	C	447	GLU	2.5
1	A	437	ILE	2.4
1	B	448	PHE	2.4
1	C	451	ILE	2.4
1	B	444[A]	LYS	2.4
1	A	500	PHE	2.4
1	C	383	PRO	2.3
1	C	293	ARG	2.3
1	B	124	GLU	2.3
1	B	81	GLY	2.2
1	C	297	ASP	2.2
1	C	495	ILE	2.2
1	B	441[A]	THR	2.1
1	C	229	ASN	2.1
1	C	470	SER	2.1
1	C	371	GLY	2.1

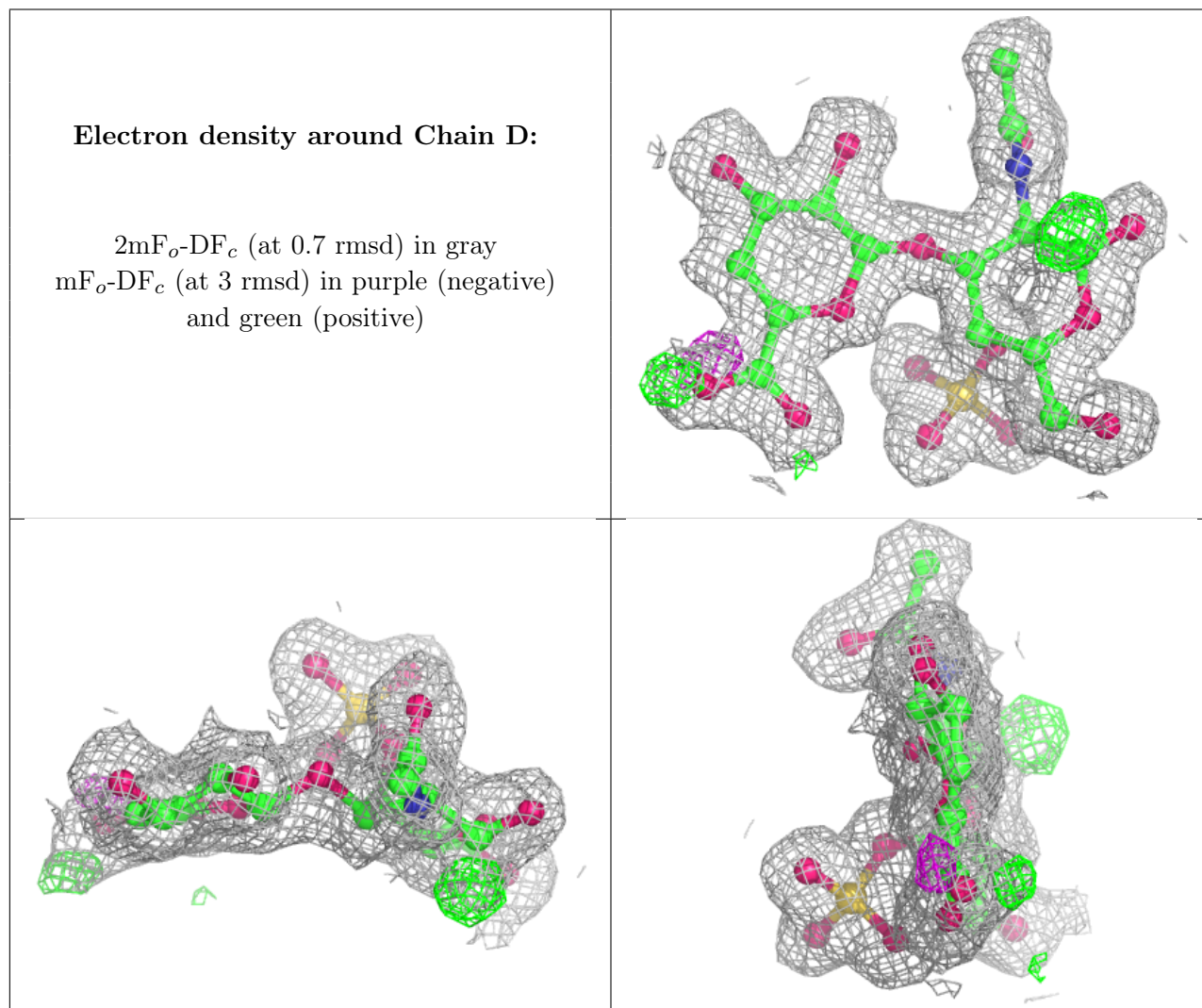## 6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates ⓘ

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.
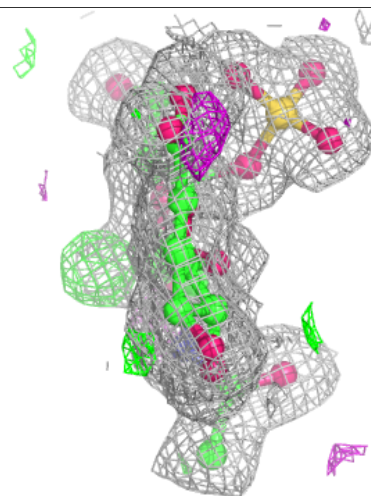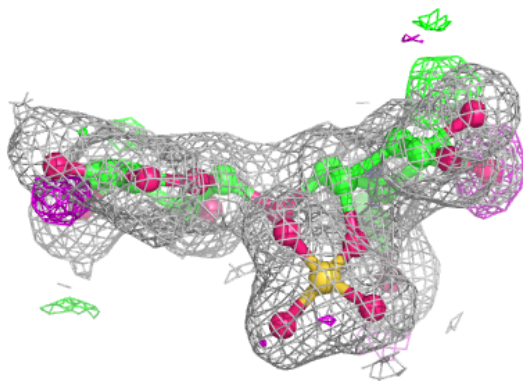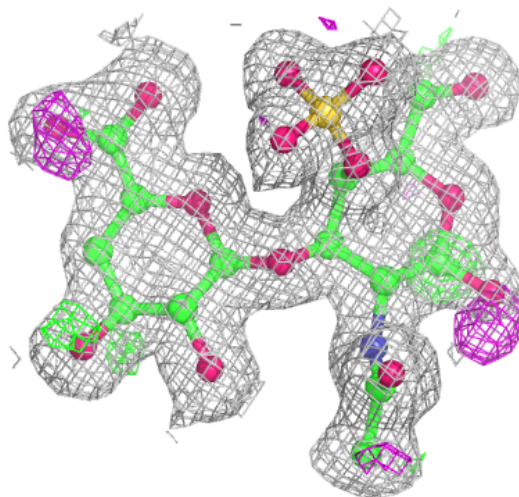
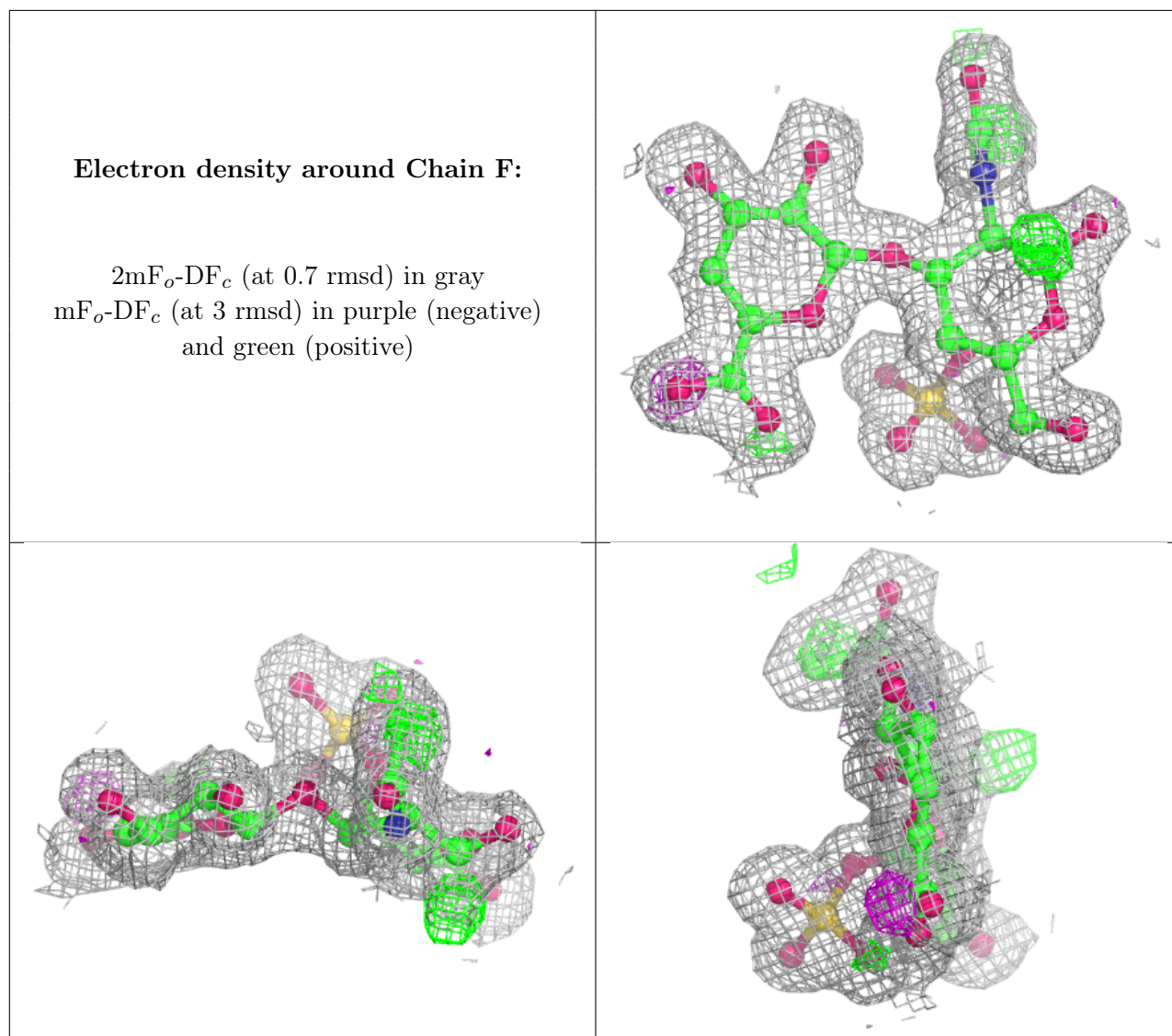| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GCD	D	2	11/12	0.94	0.10	9,11,12,13	0
2	GCD	E	2	11/12	0.94	0.11	9,11,13,14	0
2	GCD	F	2	11/12	0.94	0.11	12,13,15,15	0
2	ASG	F	1	19/19	0.95	0.12	8,11,17,22	0
2	ASG	E	1	19/19	0.97	0.10	9,11,15,19	0
2	ASG	D	1	19/19	0.97	0.10	9,11,14,19	0

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 D:

$2\text{mF}_o\text{-DF}_c$ (at 0.7 rmsd) in gray
$\text{mF}_o\text{-DF}_c$ (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)

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)

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	EDO	C	606	4/4	0.71	0.21	29,29,31,34	0
4	EDO	B	609	4/4	0.85	0.24	21,22,26,30	0
4	EDO	B	610	4/4	0.86	0.24	21,25,28,29	0
4	EDO	A	605	4/4	0.88	0.21	19,20,23,32	0
4	EDO	A	607	4/4	0.89	0.15	29,29,31,40	0
4	EDO	A	608	4/4	0.90	0.25	18,25,26,29	0
4	EDO	B	604	4/4	0.91	0.12	22,23,26,26	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EDO	C	604	4/4	0.92	0.13	23,24,28,33	0
4	EDO	B	608	4/4	0.92	0.33	28,28,30,38	0
4	EDO	B	606	4/4	0.93	0.12	20,22,23,26	0
4	EDO	B	605	4/4	0.94	0.11	24,27,29,32	0
4	EDO	B	607	4/4	0.95	0.15	13,19,22,23	0
4	EDO	C	605	4/4	0.95	0.14	19,21,21,26	0
4	EDO	A	604	4/4	0.95	0.11	15,18,20,21	0
3	CA	C	601	1/1	0.97	0.06	18,18,18,18	0
4	EDO	A	606	4/4	0.97	0.12	15,19,22,22	0
3	CA	A	601	1/1	0.99	0.05	19,19,19,19	0
3	CA	B	601	1/1	0.99	0.06	17,17,17,17	0

6.5 Other polymers ⓘ

There are no such residues in this entry.