WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 31, 2021 – 02:29 PM EDT

PDB ID	:	3H6Z
Title	:	Crystal Structure of the Four MBT Repeats of Drosophila melanogaster Sfmbt in Complex with Peptide RHR (me)K VLR
Authors	:	Grimm, C.; Mueller, C.W.
Deposited on	:	2009-04-24
Resolution	:	2.80 Å(reported)

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

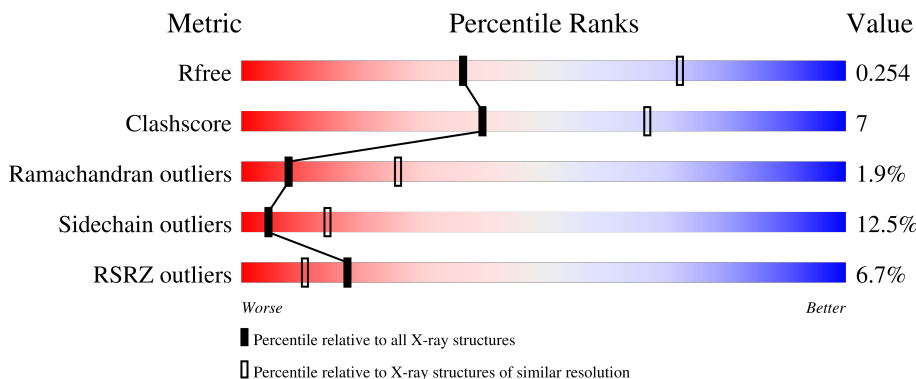|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.80 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.
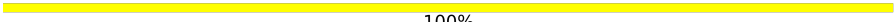
Metric	Percentile Ranks	Value
Rfree		0.254
Clashscore		7
Ramachandran outliers		1.9%
Sidechain outliers		12.5%
RSRZ outliers		6.7%

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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	447	6% / 74% 19% · ·
1	B	447	7% / 74% 18% 5% ·
2	L	6	17% / 67% 33%
2	M	6	17% / 50% 50%
3	C	2	50% 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	2	100%
3	E	2	100%
3	F	2	50% 50%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 7115 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 Polycomb protein Sfmbt.

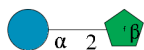| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	431	Total 3421	C 2187	N 570	O 638	S 26	0	0	0
1	B	431	Total 3421	C 2187	N 570	O 638	S 26	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	531	GLY	-	expression tag	UNP Q9VK33
A	532	ALA	-	expression tag	UNP Q9VK33
A	533	MET	-	expression tag	UNP Q9VK33
A	534	GLY	-	expression tag	UNP Q9VK33
A	715	ASP	LYS	engineered mutation	UNP Q9VK33
A	886	SER	ARG	engineered mutation	UNP Q9VK33
A	900	ASP	ARG	engineered mutation	UNP Q9VK33
B	531	GLY	-	expression tag	UNP Q9VK33
B	532	ALA	-	expression tag	UNP Q9VK33
B	533	MET	-	expression tag	UNP Q9VK33
B	534	GLY	-	expression tag	UNP Q9VK33
B	715	ASP	LYS	engineered mutation	UNP Q9VK33
B	886	SER	ARG	engineered mutation	UNP Q9VK33
B	900	ASP	ARG	engineered mutation	UNP Q9VK33

- Molecule 2 is a protein called 'HR(MLZ)VLR.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	L	6	Total 58	C 36	N 15	O 7	0	0	0
2	M	6	Total 58	C 36	N 15	O 7	0	0	0

- Molecule 3 is an oligosaccharide called beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
3	C	2	Total 23	C 12	O 11	0	0	0
3	D	2	Total 23	C 12	O 11	0	0	0
3	E	2	Total 23	C 12	O 11	0	0	0
3	F	2	Total 23	C 12	O 11	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	28	Total 28	O 28	0	0
4	L	1	Total 1	O 1	0	0
4	B	35	Total 35	O 35	0	0
4	M	1	Total 1	O 1	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 summari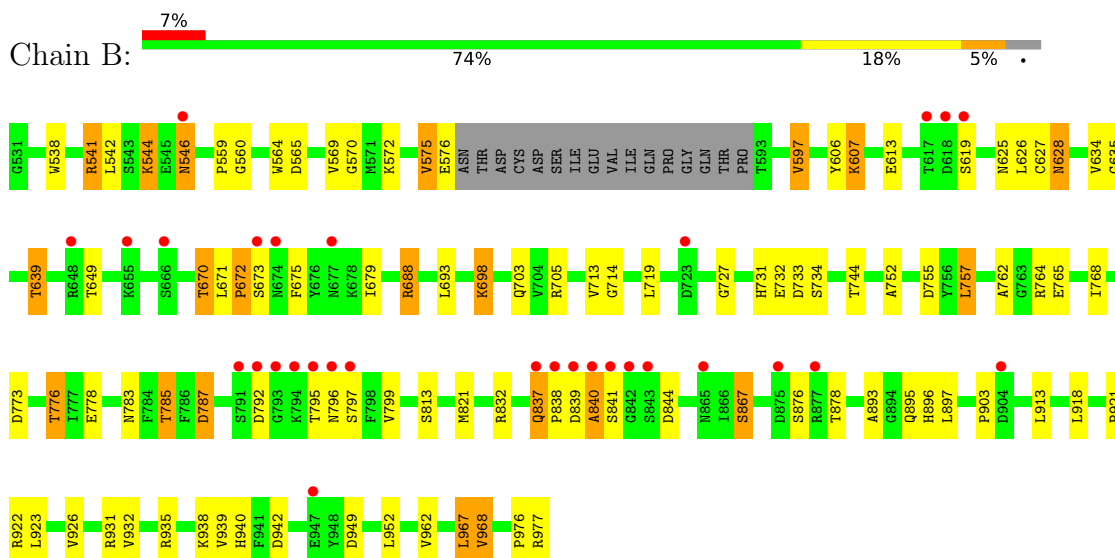ses the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: 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: Polycomb protein Sfmbt
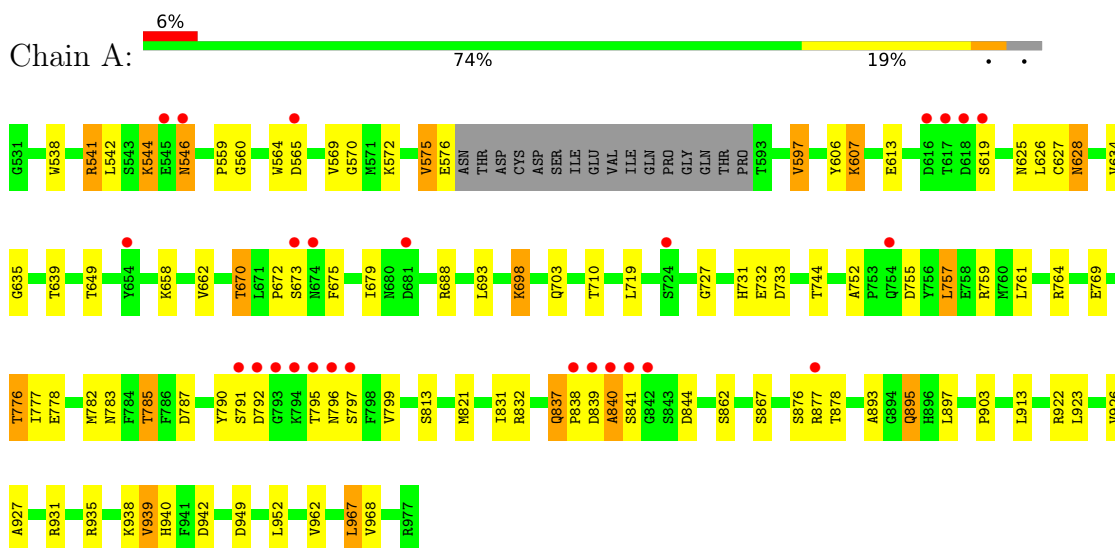
Chain A:

• Molecule 1: Polycomb protein Sfmbt

Chain B:

• Molecule 2: 'HR(MLZ)VLR

Chain L:

H18 R19 K20 V21 L22 R23

• Molecule 2: 'HR(MLZ)VLR

Chain M:

17%

50% 50%

H18 R19 K20 R23

• Molecule 3: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain C:

50% 50%

GLC1 FRU2

• Molecule 3: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain D:

100%

GLC1 FRU2

• Molecule 3: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain E:

100%

GLC1 FRU2

• Molecule 3: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain F:

50% 50%

GLC1 FRU2
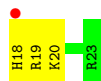
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	70.80Å 97.04Å 214.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.46 – 2.80 29.46 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.46-2.80) 99.9 (29.46-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	2.60 (at 2.80Å)	Xtriage
Refinement program	REFMAC 5.4.0065	Depositor
R, R_{free}	0.221 , 0.254 0.224 , 0.254	Depositor DCC
R_{free} test set	1884 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å²)	38.5	Xtriage
Anisotropy	0.170	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.37 , 40.6	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	7115	wwPDB-VP
Average B, all atoms (Å²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.18 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.0037e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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: MLZ, FRU, GLC

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	0.47	0/3521	0.63	0/4791
1	B	0.47	0/3521	0.63	0/4791
2	L	0.95	0/47	0.80	0/58
2	M	0.90	0/47	0.79	0/58
All	All	0.48	0/7136	0.63	0/9698

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	3421	0	3255	49	0
1	B	3421	0	3255	53	0
2	L	58	0	66	0	0
2	M	58	0	66	0	0
3	C	23	0	21	2	0
3	D	23	0	21	0	0
3	E	23	0	21	3	0
3	F	23	0	21	0	0
4	A	28	0	0	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	35	0	0	7	0
4	L	1	0	0	0	0
4	M	1	0	0	0	0
All	All	7115	0	6726	101	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 7.

All (101) 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:634:VAL:H	1:B:670:THR:HG21	1.21	1.00
1:A:634:VAL:H	1:A:670:THR:HG21	1.27	0.98
1:A:542:LEU:HD11	1:A:607:LYS:HD3	1.54	0.90
1:A:797:SER:HB3	1:A:897:LEU:HD22	1.53	0.90
1:B:542:LEU:HD11	1:B:607:LYS:HD3	1.52	0.89
1:B:572:LYS:O	1:B:670:THR:HG23	1.75	0.85
1:B:698:LYS:HB2	3:E:2:FRU:H12	1.59	0.84
1:A:731:HIS:HD2	1:A:733:ASP:H	1.25	0.83
1:B:731:HIS:HD2	1:B:733:ASP:H	1.24	0.82
1:B:867:SER:HB2	4:B:50:HOH:O	1.77	0.82
1:A:877:ARG:HB2	4:A:46:HOH:O	1.79	0.81
1:B:797:SER:HB3	1:B:897:LEU:HD22	1.60	0.81
1:B:785:THR:HB	4:B:19:HOH:O	1.82	0.78
1:A:903:PRO:HD2	1:A:935:ARG:HG3	1.65	0.78
1:A:572:LYS:O	1:A:670:THR:HG23	1.85	0.77
1:B:903:PRO:HD2	1:B:935:ARG:HG3	1.68	0.76
1:B:634:VAL:N	1:B:670:THR:HG21	2.02	0.71
1:A:634:VAL:N	1:A:670:THR:HG21	2.05	0.68
1:B:625:ASN:HB3	1:B:628:ASN:ND2	2.09	0.67
1:A:625:ASN:HB3	1:A:628:ASN:HD21	1.60	0.67
1:B:635:GLY:O	1:B:639:THR:HG22	1.93	0.67
1:B:625:ASN:HB3	1:B:628:ASN:HD21	1.59	0.67
1:A:625:ASN:HB3	1:A:628:ASN:ND2	2.10	0.66
1:A:575:VAL:HG12	1:A:576:GLU:H	1.60	0.66
1:B:693:LEU:HD11	1:B:719:LEU:HD22	1.78	0.64
1:A:575:VAL:O	1:A:576:GLU:HB2	1.98	0.63
1:B:731:HIS:CD2	1:B:733:ASP:H	2.12	0.62
1:B:575:VAL:HG12	1:B:576:GLU:H	1.65	0.62
1:B:796:ASN:OD1	4:B:62:HOH:O	2.16	0.61
1:A:635:GLY:O	1:A:639:THR:HG22	2.00	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:693:LEU:HD11	1:A:719:LEU:HD22	1.83	0.59
1:A:698:LYS:HB2	3:C:2:FRU:H12	1.85	0.58
1:B:837:GLN:HG2	1:B:838:PRO:HD2	1.86	0.57
1:B:575:VAL:O	1:B:576:GLU:HB2	2.04	0.57
1:A:837:GLN:HG2	1:A:838:PRO:HD2	1.88	0.56
1:A:731:HIS:CD2	1:A:733:ASP:H	2.14	0.56
1:A:752:ALA:HB3	1:A:757:LEU:HD13	1.89	0.55
1:B:597:VAL:HG23	1:B:613:GLU:HG2	1.90	0.54
1:B:572:LYS:O	1:B:670:THR:CG2	2.52	0.53
1:A:710:THR:HG21	4:A:39:HOH:O	2.08	0.53
1:B:776:THR:HB	1:B:778:GLU:OE1	2.10	0.52
1:A:785:THR:HG22	4:A:47:HOH:O	2.11	0.51
1:A:675:PHE:O	1:A:679:ILE:HG12	2.11	0.50
1:B:813:SER:HB3	1:B:949:ASP:O	2.12	0.50
1:A:572:LYS:O	1:A:670:THR:CG2	2.59	0.49
1:B:559:PRO:HD2	1:B:606:TYR:OH	2.12	0.49
1:A:628:ASN:N	1:A:628:ASN:HD22	2.11	0.49
1:A:559:PRO:HD2	1:A:606:TYR:OH	2.12	0.48
1:A:776:THR:HB	1:A:778:GLU:OE1	2.13	0.48
1:B:731:HIS:HD2	1:B:733:ASP:N	2.02	0.48
1:B:821:MET:HG3	1:B:832:ARG:HB2	1.96	0.48
1:A:560:GLY:O	1:A:564:TRP:HB2	2.14	0.47
1:A:719:LEU:O	1:A:727:GLY:HA3	2.14	0.47
1:B:628:ASN:HD22	1:B:628:ASN:N	2.12	0.47
1:B:719:LEU:O	1:B:727:GLY:HA3	2.14	0.47
1:B:752:ALA:HB3	1:B:757:LEU:HD13	1.97	0.47
1:B:785:THR:CB	4:B:19:HOH:O	2.53	0.47
1:A:931:ARG:HB2	1:A:938:LYS:HB3	1.95	0.47
1:A:627:CYS:SG	1:A:967:LEU:HG	2.55	0.47
1:B:560:GLY:O	1:B:564:TRP:HB2	2.15	0.46
1:B:714:GLY:HA3	1:B:968:VAL:HG22	1.97	0.46
1:A:796:ASN:OD1	4:A:27:HOH:O	2.21	0.46
1:B:940:HIS:HD2	1:B:949:ASP:OD1	1.98	0.46
1:B:734:SER:HB3	4:B:61:HOH:O	2.14	0.46
1:A:597:VAL:HG23	1:A:613:GLU:HG2	1.98	0.46
1:B:931:ARG:HB2	1:B:938:LYS:HB3	1.97	0.46
1:A:840:ALA:CB	1:A:844:ASP:HB2	2.45	0.46
1:A:877:ARG:CB	4:A:46:HOH:O	2.54	0.45
1:A:790:TYR:HB2	1:B:762:ALA:O	2.16	0.45
1:B:675:PHE:O	1:B:679:ILE:HG12	2.17	0.44
1:B:840:ALA:CB	1:B:844:ASP:HB2	2.46	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:895:GLN:HE21	1:A:895:GLN:HB3	1.60	0.44
1:B:893:ALA:HB1	1:B:897:LEU:HD12	1.98	0.44
1:B:597:VAL:CG2	1:B:613:GLU:CG	2.96	0.44
1:B:627:CYS:SG	1:B:967:LEU:HG	2.58	0.44
1:B:541:ARG:O	1:B:544:LYS:HG3	2.18	0.43
1:B:597:VAL:CG2	1:B:613:GLU:HG2	2.48	0.43
1:A:538:TRP:CH2	1:A:926:VAL:HG13	2.53	0.43
1:B:688:ARG:NH2	1:B:773:ASP:OD1	2.52	0.43
1:A:541:ARG:O	1:A:544:LYS:HG3	2.19	0.43
1:A:821:MET:HG3	1:A:832:ARG:HB2	2.01	0.43
1:A:813:SER:HB3	1:A:949:ASP:O	2.19	0.43
1:A:927:ALA:HB1	1:A:939:VAL:HG23	2.00	0.43
1:A:940:HIS:HD2	1:A:949:ASP:OD1	2.02	0.42
1:B:787:ASP:HB2	4:B:19:HOH:O	2.20	0.42
1:B:538:TRP:CH2	1:B:926:VAL:HG13	2.54	0.42
1:B:671:LEU:HA	1:B:672:PRO:HD2	1.89	0.42
1:A:893:ALA:HB1	1:A:897:LEU:HD12	2.01	0.42
1:A:744:THR:HG23	4:A:978:HOH:O	2.19	0.42
1:A:597:VAL:CG2	1:A:613:GLU:CG	2.98	0.41
1:A:658:LYS:O	1:A:662:VAL:HG23	2.20	0.41
1:B:744:THR:HG23	4:B:57:HOH:O	2.19	0.41
1:A:777:ILE:HD13	1:B:768:ILE:HD12	2.01	0.41
1:B:714:GLY:HA3	1:B:968:VAL:CG2	2.50	0.41
1:A:831:ILE:HD13	1:A:831:ILE:HA	1.90	0.41
1:B:896:HIS:CD2	1:B:896:HIS:H	2.37	0.41
1:A:698:LYS:HB2	3:C:2:FRU:C1	2.49	0.41
1:B:921:PRO:HD3	3:E:2:FRU:H62	2.02	0.41
1:A:777:ILE:HD11	1:A:782:MET:SD	2.61	0.41
3:E:1:GLC:C6	3:E:2:FRU:H61	2.51	0.41
1:B:976:PRO:O	1:B:977:ARG:HG2	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	427/447 (96%)	393 (92%)	26 (6%)	8 (2%)	8	26
1	B	427/447 (96%)	396 (93%)	23 (5%)	8 (2%)	8	26
2	L	3/6 (50%)	2 (67%)	1 (33%)	0	100	100
2	M	3/6 (50%)	2 (67%)	1 (33%)	0	100	100
All	All	860/906 (95%)	793 (92%)	51 (6%)	16 (2%)	8	26

All (16) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	841	SER
1	B	841	SER
1	B	732	GLU
1	B	840	ALA
1	A	546	ASN
1	A	672	PRO
1	A	732	GLU
1	A	840	ALA
1	B	546	ASN
1	B	672	PRO
1	A	839	ASP
1	B	839	ASP
1	A	569	VAL
1	B	570	GLY
1	A	570	GLY
1	B	569	VAL

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	371/386 (96%)	326 (88%)	45 (12%)	5	15
1	B	371/386 (96%)	325 (88%)	46 (12%)	4	14

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	L	5/5 (100%)	4 (80%)	1 (20%)	1	4
2	M	5/5 (100%)	3 (60%)	2 (40%)	0	0
All	All	752/782 (96%)	658 (88%)	94 (12%)	4	14

All (94) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	541	ARG
1	A	544	LYS
1	A	546	ASN
1	A	565	ASP
1	A	575	VAL
1	A	597	VAL
1	A	607	LYS
1	A	619	SER
1	A	626	LEU
1	A	628	ASN
1	A	649	THR
1	A	670	THR
1	A	673	SER
1	A	688	ARG
1	A	698	LYS
1	A	703	GLN
1	A	755	ASP
1	A	757	LEU
1	A	759	ARG
1	A	761	LEU
1	A	764	ARG
1	A	769	GLU
1	A	776	THR
1	A	783	ASN
1	A	785	THR
1	A	787	ASP
1	A	791	SER
1	A	792	ASP
1	A	795	THR
1	A	799	VAL
1	A	837	GLN
1	A	862	SER
1	A	867	SER
1	A	876	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	878	THR
1	A	895	GLN
1	A	913	LEU
1	A	922	ARG
1	A	923	LEU
1	A	939	VAL
1	A	942	ASP
1	A	952	LEU
1	A	962	VAL
1	A	967	LEU
1	A	968	VAL
2	L	21	VAL
1	B	541	ARG
1	B	544	LYS
1	B	546	ASN
1	B	565	ASP
1	B	575	VAL
1	B	597	VAL
1	B	607	LYS
1	B	619	SER
1	B	626	LEU
1	B	628	ASN
1	B	639	THR
1	B	649	THR
1	B	670	THR
1	B	673	SER
1	B	688	ARG
1	B	698	LYS
1	B	703	GLN
1	B	705	ARG
1	B	713	VAL
1	B	755	ASP
1	B	757	LEU
1	B	764	ARG
1	B	765	GLU
1	B	776	THR
1	B	783	ASN
1	B	785	THR
1	B	787	ASP
1	B	792	ASP
1	B	795	THR
1	B	799	VAL

Continued from previous page...

Mol	Chain	Res	Type
1	B	837	GLN
1	B	867	SER
1	B	876	SER
1	B	878	THR
1	B	895	GLN
1	B	913	LEU
1	B	918	LEU
1	B	922	ARG
1	B	923	LEU
1	B	932	VAL
1	B	939	VAL
1	B	942	ASP
1	B	952	LEU
1	B	962	VAL
1	B	967	LEU
1	B	968	VAL
2	M	18	HIS
2	M	19	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	628	ASN
1	A	692	ASN
1	A	731	HIS
1	A	837	GLN
1	A	895	GLN
1	A	896	HIS
1	A	940	HIS
1	B	628	ASN
1	B	692	ASN
1	B	731	HIS
1	B	837	GLN
1	B	895	GLN
1	B	896	HIS
1	B	940	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	MLZ	M	20	2	8,9,10	0.81	0	4,9,11	1.34	1 (25%)
2	MLZ	L	20	2	8,9,10	0.76	0	4,9,11	1.53	1 (25%)

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	MLZ	M	20	2	-	1/7/8/10	-
2	MLZ	L	20	2	-	2/7/8/10	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	L	20	MLZ	CM-NZ-CE	2.52	119.25	111.95
2	M	20	MLZ	CM-NZ-CE	2.22	118.36	111.95

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	L	20	MLZ	C-CA-CB-CG
2	M	20	MLZ	C-CA-CB-CG
2	L	20	MLZ	CE-CD-CG-CB

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

8 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	GLC	C	1	3	11,11,12	0.29	0	15,15,17	1.77	4 (26%)
3	FRU	C	2	3	11,12,12	0.84	1 (9%)	10,18,18	1.05	1 (10%)
3	GLC	D	1	3	11,11,12	0.33	0	15,15,17	1.22	1 (6%)
3	FRU	D	2	3	11,12,12	0.82	1 (9%)	10,18,18	0.88	0
3	GLC	E	1	3	11,11,12	0.27	0	15,15,17	1.84	4 (26%)
3	FRU	E	2	3	11,12,12	0.88	1 (9%)	10,18,18	0.97	1 (10%)
3	GLC	F	1	3	11,11,12	0.40	0	15,15,17	1.38	3 (20%)
3	FRU	F	2	3	11,12,12	0.76	0	10,18,18	0.73	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
3	GLC	C	1	3	-	2/2/19/22	0/1/1/1
3	FRU	C	2	3	-	5/5/24/24	0/1/1/1
3	GLC	D	1	3	-	0/2/19/22	0/1/1/1
3	FRU	D	2	3	-	4/5/24/24	0/1/1/1
3	GLC	E	1	3	-	2/2/19/22	0/1/1/1
3	FRU	E	2	3	-	5/5/24/24	0/1/1/1
3	GLC	F	1	3	-	2/2/19/22	0/1/1/1
3	FRU	F	2	3	-	1/5/24/24	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	2	FRU	O2-C2	2.65	1.45	1.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	2	FRU	O2-C2	2.42	1.44	1.40
3	D	2	FRU	O2-C2	2.10	1.44	1.40

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	1	GLC	O5-C1-C2	-4.34	104.07	110.77
3	D	1	GLC	C1-O5-C5	3.91	117.50	112.19
3	C	1	GLC	O5-C1-C2	-3.81	104.89	110.77
3	C	1	GLC	C3-C4-C5	3.02	115.63	110.24
3	E	1	GLC	C3-C4-C5	2.90	115.41	110.24
3	F	1	GLC	O5-C1-C2	-2.70	106.60	110.77
3	F	1	GLC	O5-C5-C6	2.53	111.17	107.20
3	C	2	FRU	O2-C2-O5	-2.42	104.84	109.50
3	E	1	GLC	O4-C4-C5	-2.33	103.50	109.30
3	E	1	GLC	O4-C4-C3	-2.19	105.29	110.35
3	E	2	FRU	O2-C2-O5	-2.16	105.34	109.50
3	F	1	GLC	C1-C2-C3	-2.13	107.05	109.67
3	C	1	GLC	O5-C5-C6	2.11	110.51	107.20
3	C	1	GLC	O4-C4-C5	-2.01	104.31	109.30

There are no chirality outliers.

All (21) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	2	FRU	O1-C1-C2-C3
3	C	2	FRU	O1-C1-C2-O2
3	C	2	FRU	O5-C5-C6-O6
3	E	2	FRU	O1-C1-C2-C3
3	E	2	FRU	O1-C1-C2-O2
3	E	2	FRU	O5-C5-C6-O6
3	C	1	GLC	O5-C5-C6-O6
3	E	1	GLC	O5-C5-C6-O6
3	C	2	FRU	C4-C5-C6-O6
3	E	2	FRU	C4-C5-C6-O6
3	C	1	GLC	C4-C5-C6-O6
3	E	1	GLC	C4-C5-C6-O6
3	F	1	GLC	C4-C5-C6-O6
3	F	1	GLC	O5-C5-C6-O6
3	E	2	FRU	O1-C1-C2-O5
3	C	2	FRU	O1-C1-C2-O5
3	D	2	FRU	O1-C1-C2-O2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	D	2	FRU	O1-C1-C2-O5
3	D	2	FRU	O1-C1-C2-C3
3	D	2	FRU	C4-C5-C6-O6
3	F	2	FRU	O5-C5-C6-O6

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	2	FRU	2	0
3	E	1	GLC	1	0
3	E	2	FRU	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 oligosaccharide.

Oligosaccharide Chain C

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain D

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain F

Bond lengths

Bond angles

Torsions

Rings
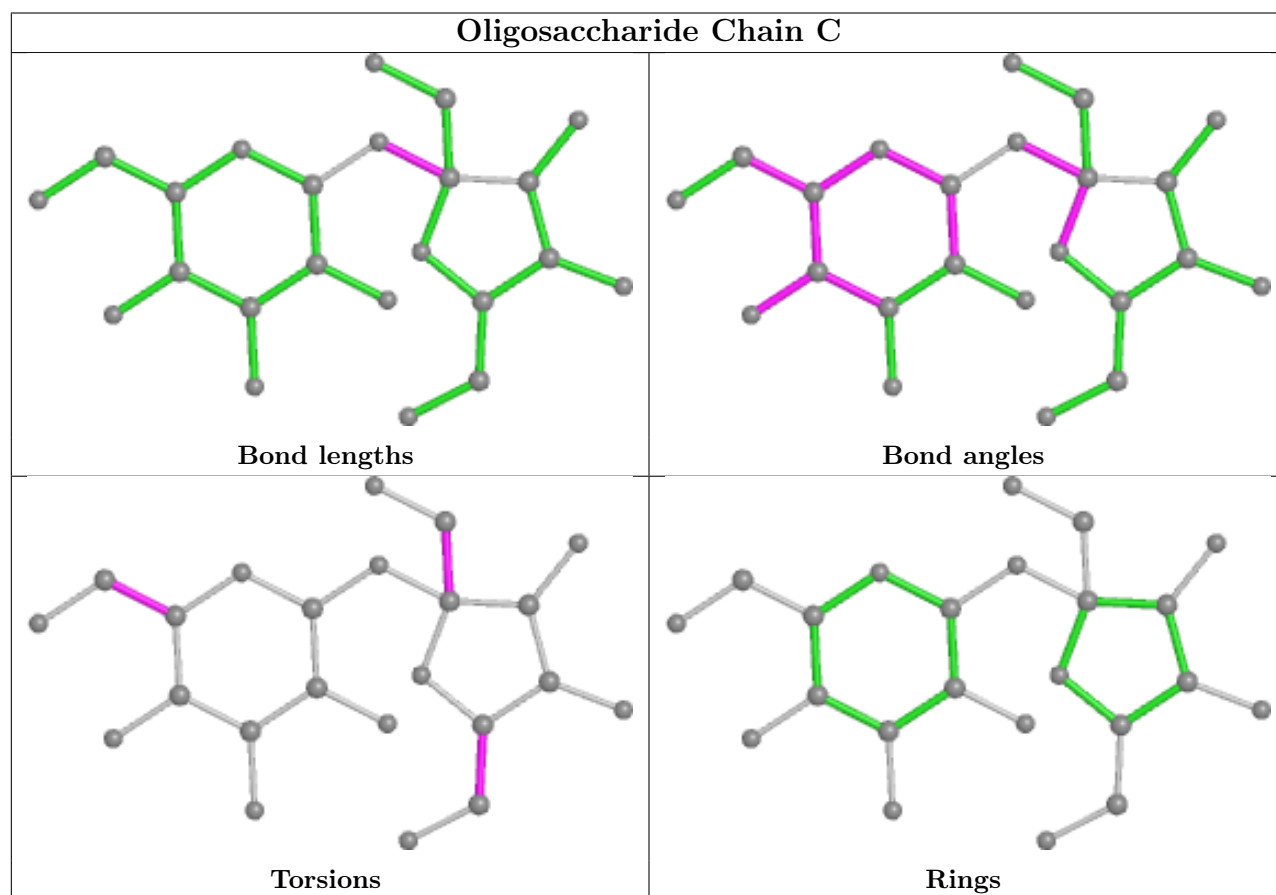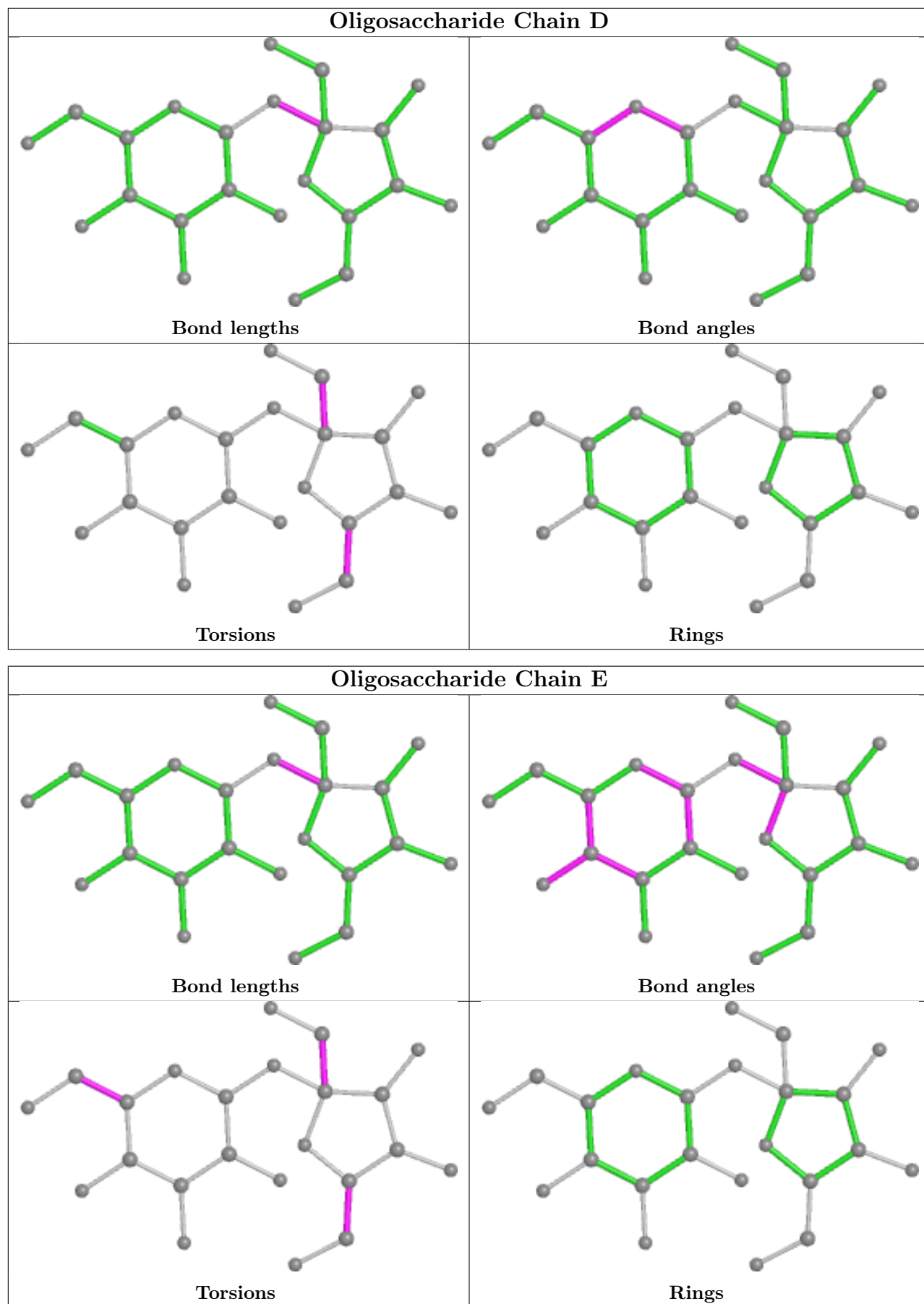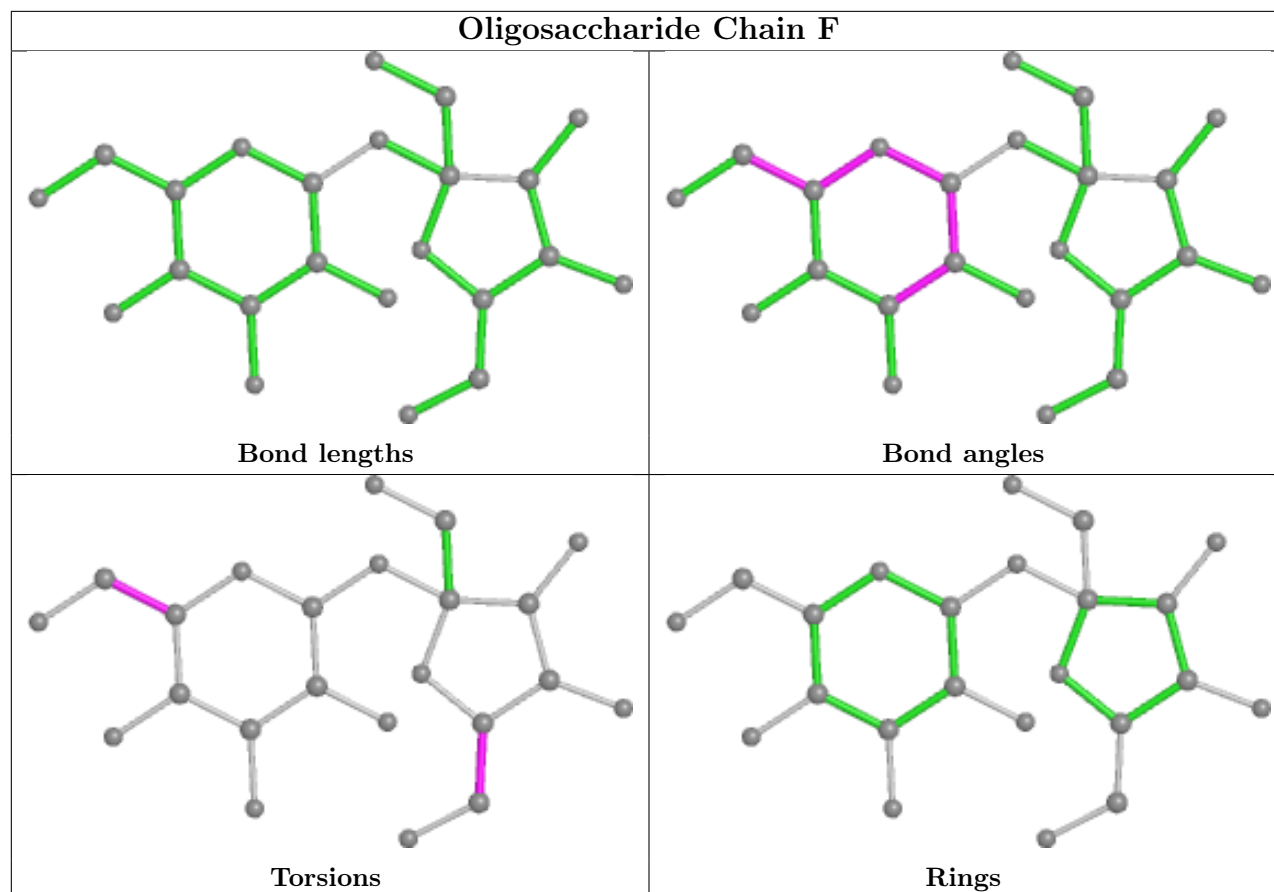
5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

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	431/447 (96%)	0.10	26 (6%)	21	14	21, 33, 65, 90	0
1	B	431/447 (96%)	0.08	30 (6%)	16	9	20, 34, 64, 86	0
2	L	5/6 (83%)	1.86	1 (20%)	1	0	51, 51, 63, 69	0
2	M	5/6 (83%)	1.67	1 (20%)	1	0	51, 51, 63, 69	0
All	All	872/906 (96%)	0.11	58 (6%)	17	10	20, 34, 65, 90	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	L	18	HIS	9.5
2	M	18	HIS	7.7
1	B	795	THR	7.4
1	A	795	THR	5.4
1	A	673	SER	5.3
1	B	839	ASP	5.0
1	A	877	ARG	4.8
1	B	673	SER	4.8
1	A	840	ALA	4.4
1	B	838	PRO	4.2
1	B	877	ARG	4.0
1	A	838	PRO	3.8
1	B	842	GLY	3.8
1	A	796	ASN	3.7
1	B	837	GLN	3.7
1	A	681	ASP	3.6
1	B	617	THR	3.5
1	B	904	ASP	3.4
1	B	793	GLY	3.4
1	A	793	GLY	3.3
1	A	792	ASP	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	618	ASP	3.3
1	A	839	ASP	3.3
1	A	546	ASN	3.2
1	B	796	ASN	3.2
1	B	841	SER	3.1
1	B	791	SER	3.0
1	B	840	ALA	3.0
1	A	754	GLN	3.0
1	B	843	SER	2.9
1	B	619	SER	2.8
1	A	565	ASP	2.8
1	A	841	SER	2.8
1	B	792	ASP	2.7
1	B	723	ASP	2.6
1	B	648	ARG	2.6
1	A	797	SER	2.5
1	A	794	LYS	2.5
1	A	616	ASP	2.5
1	A	617	THR	2.5
1	A	791	SER	2.5
1	B	655	LYS	2.5
1	A	724	SER	2.4
1	A	674	ASN	2.4
1	B	797	SER	2.4
1	A	619	SER	2.4
1	B	677	ASN	2.4
1	B	947	GLU	2.3
1	B	666	SER	2.3
1	A	842	GLY	2.3
1	B	794	LYS	2.3
1	B	546	ASN	2.3
1	A	618	ASP	2.3
1	B	875	ASP	2.2
1	B	674	ASN	2.2
1	A	654	TYR	2.1
1	A	545	GLU	2.0
1	B	865	ASN	2.0

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

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	MLZ	L	20	10/11	0.93	0.17	42,51,55,58	0
2	MLZ	M	20	10/11	0.94	0.20	42,51,55,58	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	GLC	F	1	11/12	0.80	0.25	91,91,93,94	0
3	FRU	F	2	12/12	0.84	0.32	91,91,92,92	0
3	GLC	E	1	11/12	0.89	0.21	44,45,47,49	0
3	FRU	C	2	12/12	0.91	0.20	41,45,46,49	0
3	GLC	C	1	11/12	0.91	0.18	44,45,47,49	0
3	FRU	D	2	12/12	0.92	0.24	56,59,59,60	0
3	GLC	D	1	11/12	0.92	0.25	58,59,60,61	0
3	FRU	E	2	12/12	0.93	0.20	41,45,46,49	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.
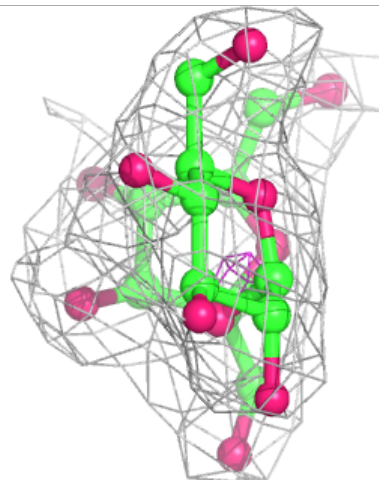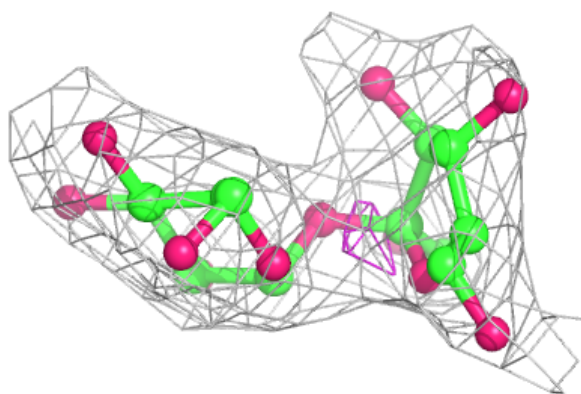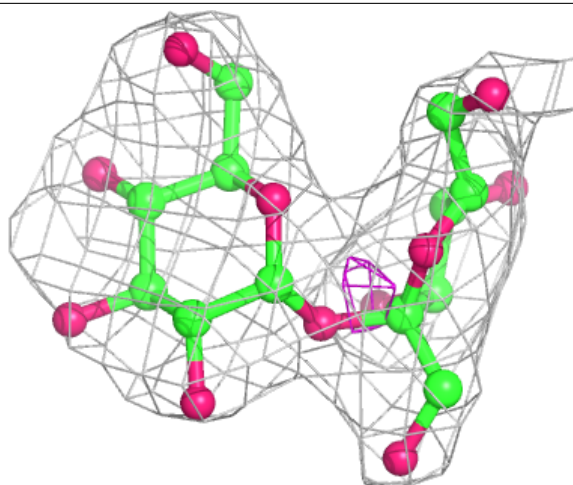
Electron density around Chain C:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
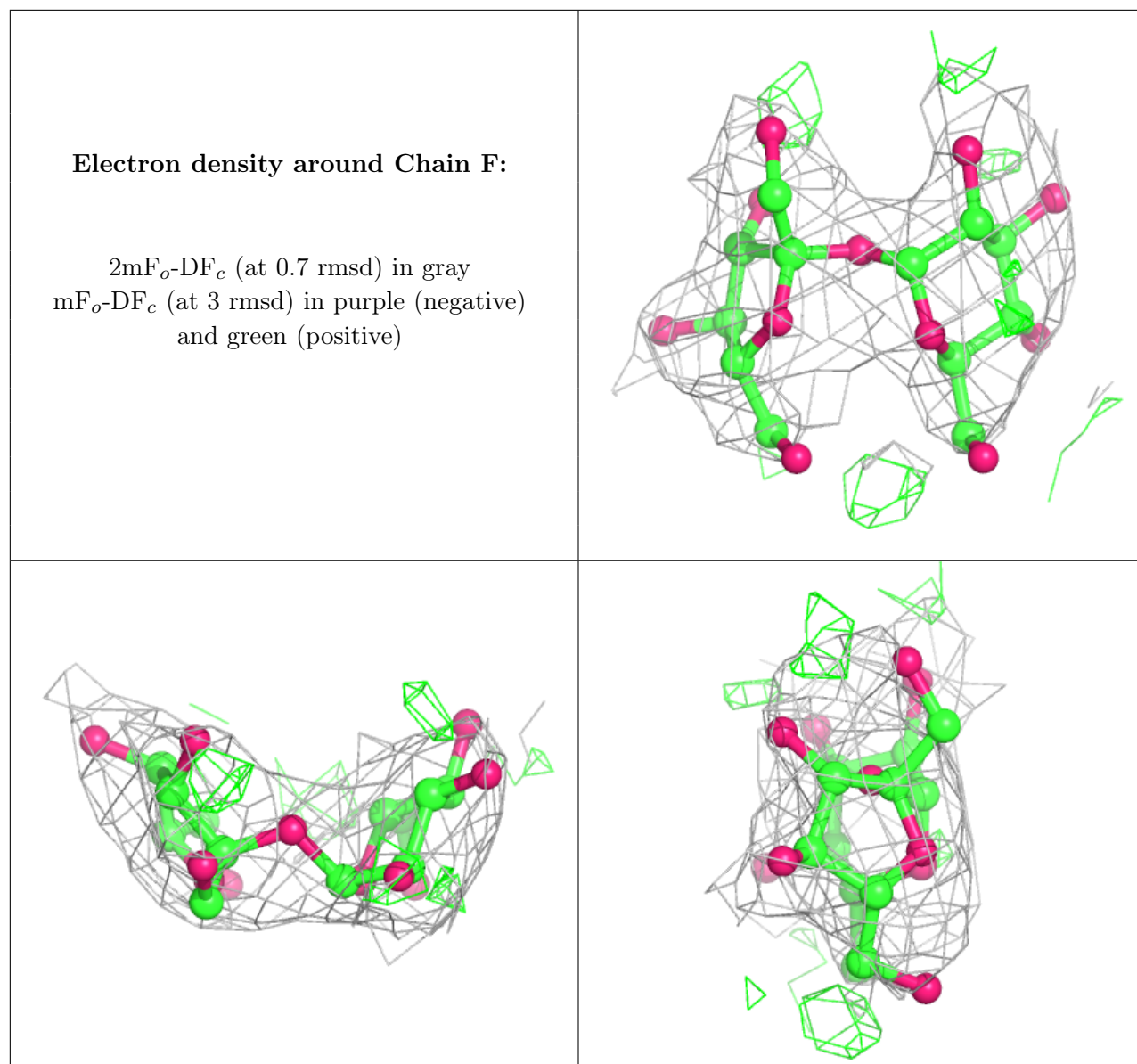and green (positive)
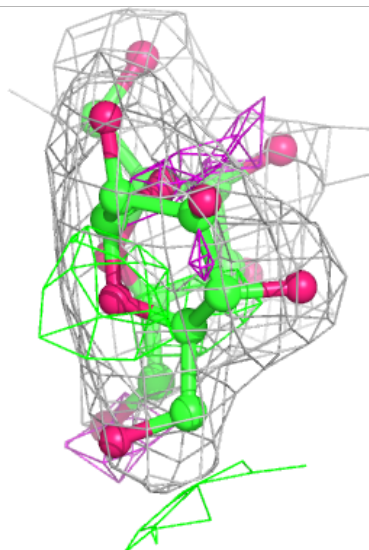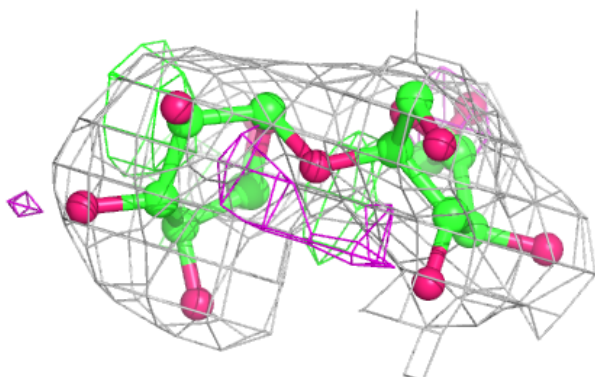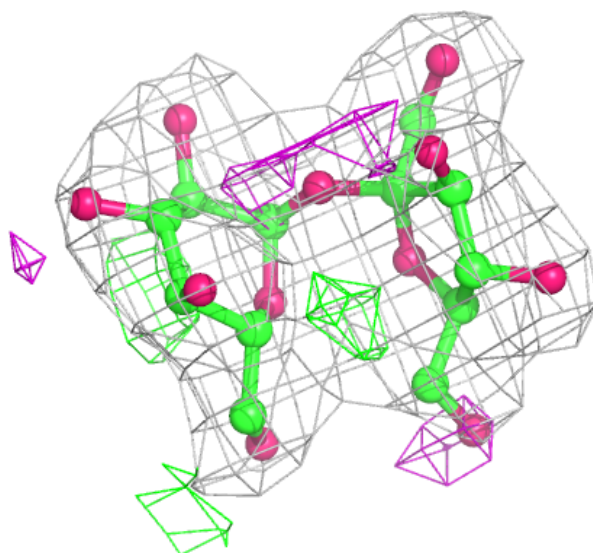
Electron density around Chain D:

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)

Electron density around Chain F:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
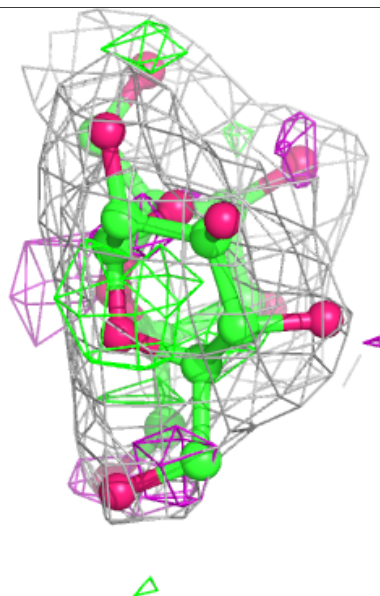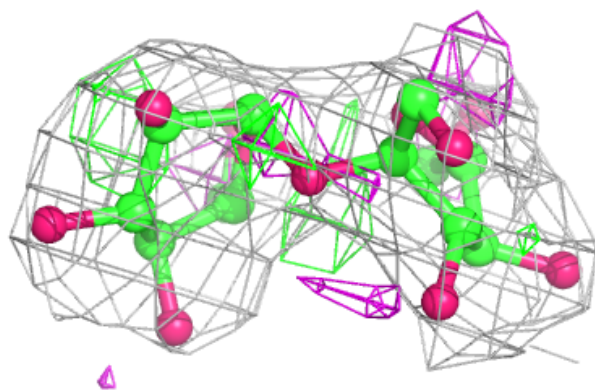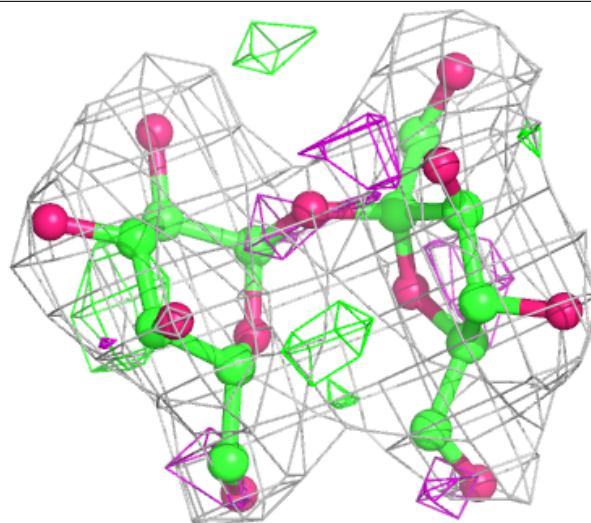
6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.