Full wwPDB X-ray Structure Validation Report (i)

Sep 26, 2023 – 04:53 PM EDT

PDB ID	:	6CF5
Title	:	Crystal structure of the A/Viet Nam/1203/2004(H5N1) influenza virus hemagglutinin in complex with small molecule N-Cyclohexyltaurine
Authors	:	Wilson, I.A.; Kadam, R.U.
Deposited on	:	2018-02-13
Resolution	:	2.04 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:
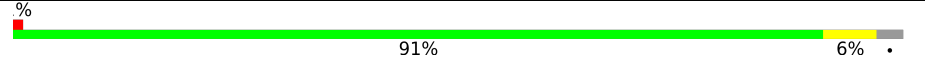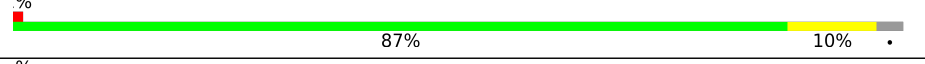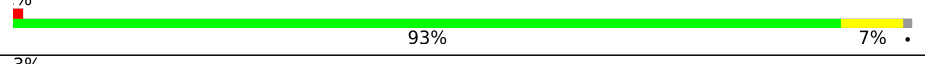*X-RAY DIFFRACTION*
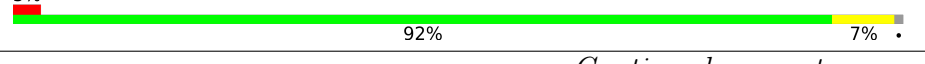
The reported resolution of this entry is 2.04 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
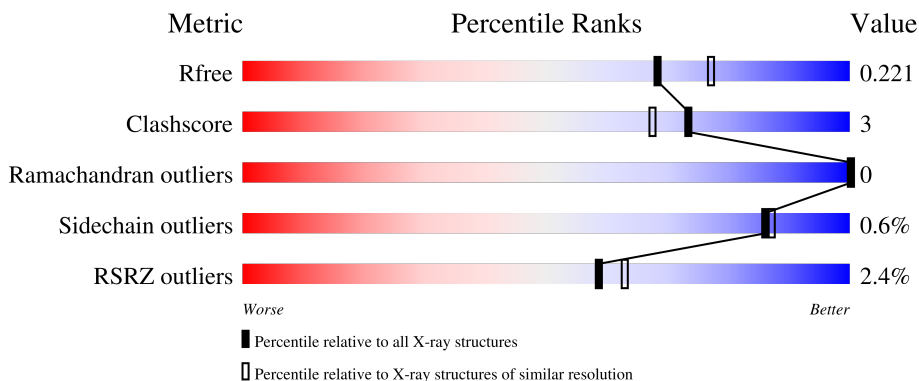
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	334	% / 91% / 6% ·
1	C	334	% / 87% / 10% ·
1	E	334	% / 87% / 10% ·
2	B	177	% / 93% / 7% ·
2	D	177	3% / 92% / 7% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	177	10% ... 94% ... 5%
3	G	2	100%
3	H	2	100%
3	I	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
3	NAG	G	2	-	-	-	X
5	NAG	C	402	-	-	-	X

2 Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 13578 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 Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	324	Total 2579	C 1629	N 446	O 489	S 15	0	2	0
1	C	323	Total 2560	C 1618	N 442	O 485	S 15	0	0	0
1	E	323	Total 2576	C 1630	N 444	O 486	S 16	0	3	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	expression tag	UNP Q5EP31
A	8	ASP	-	expression tag	UNP Q5EP31
A	9	PRO	-	expression tag	UNP Q5EP31
A	10	GLY	-	expression tag	UNP Q5EP31
C	7	ALA	-	expression tag	UNP Q5EP31
C	8	ASP	-	expression tag	UNP Q5EP31
C	9	PRO	-	expression tag	UNP Q5EP31
C	10	GLY	-	expression tag	UNP Q5EP31
E	7	ALA	-	expression tag	UNP Q5EP31
E	8	ASP	-	expression tag	UNP Q5EP31
E	9	PRO	-	expression tag	UNP Q5EP31
E	10	GLY	-	expression tag	UNP Q5EP31

- Molecule 2 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	176	Total 1427	C 888	N 247	O 284	S 8	0	1	0
2	D	176	Total 1433	C 891	N 247	O 286	S 9	0	2	0
2	F	175	Total 1423	C 886	N 246	O 283	S 8	0	1	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	175	SER	-	expression tag	UNP Q6DQ18
B	176	GLY	-	expression tag	UNP Q6DQ18
B	177	ARG	-	expression tag	UNP Q6DQ18
D	175	SER	-	expression tag	UNP Q6DQ18
D	176	GLY	-	expression tag	UNP Q6DQ18
D	177	ARG	-	expression tag	UNP Q6DQ18
F	175	SER	-	expression tag	UNP Q6DQ18
F	176	GLY	-	expression tag	UNP Q6DQ18
F	177	ARG	-	expression tag	UNP Q6DQ18

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	G	2	Total 28	C 16	N 2	O 10	0	0	0
3	H	2	Total 28	C 16	N 2	O 10	0	0	0
3	I	2	Total 28	C 16	N 2	O 10	0	0	0

- Molecule 4 is 2-[N-CYCLOHEXYLAMINO]ETHANE SULFONIC ACID (three-letter code: NHE) (formula: $C_8H_{17}NO_3S$).
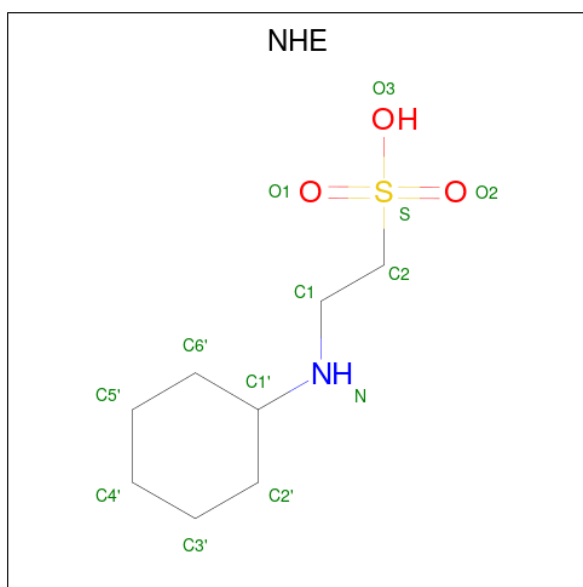
NHE

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
4	C	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
4	E	1	Total	C	N	O	S	0	0
			13	8	1	3	1		

• Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

NAG

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 14	C 8	N 1	O 5	0	0
5	C	1	Total 14	C 8	N 1	O 5	0	0
5	C	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).
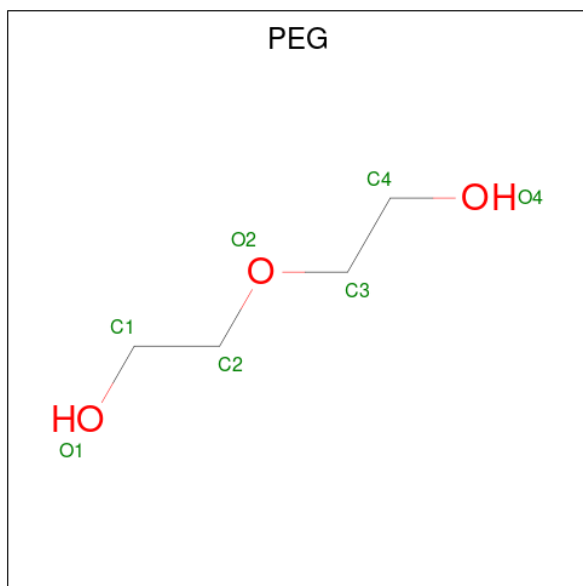
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total 7	C 4	O 3	0	0
6	A	1	Total 7	C 4	O 3	0	0
6	A	1	Total 7	C 4	O 3	0	0
6	B	1	Total 7	C 4	O 3	0	0
6	B	1	Total 7	C 4	O 3	0	0
6	C	1	Total 7	C 4	O 3	0	0
6	D	1	Total 7	C 4	O 3	0	0
6	D	1	Total 7	C 4	O 3	0	0
6	E	1	Total 7	C 4	O 3	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 6	C 3	O 3	0	0
7	C	1	Total 6	C 3	O 3	0	0
7	E	1	Total 6	C 3	O 3	0	0
7	E	1	Total 6	C 3	O 3	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	335	Total 335	O 335	0	0
8	B	124	Total 124	O 124	0	0
8	C	311	Total 311	O 311	0	0
8	D	151	Total 151	O 151	0	0
8	E	283	Total 283	O 283	0	0
8	F	124	Total 124	O 124	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: Hemagglutinin

Chain A:

- Molecule 1: Hemagglutinin

Chain C:

- Molecule 1: Hemagglutinin

Chain E:

- Molecule 2: Hemagglutinin

Chain B:

- Molecule 2: Hemagglutinin

Chain D:

- Molecule 2: Hemagglutinin

Chain F:

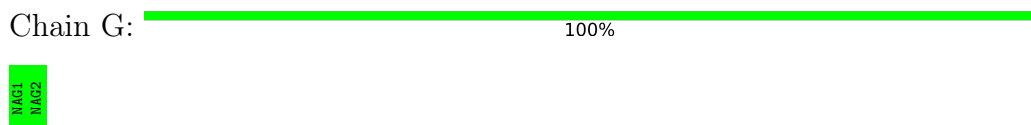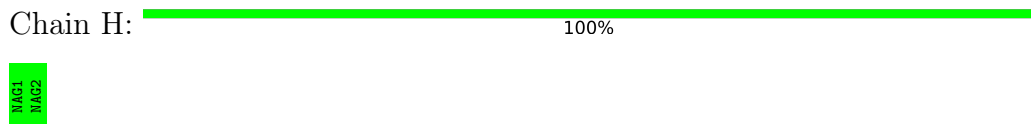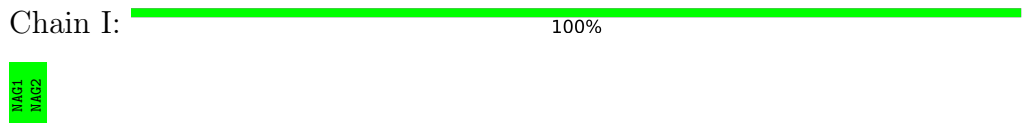- Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:

- Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain H:

- Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain I:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	72.58Å 232.98Å 72.63Å 90.00° 118.91° 90.00°	Depositor
Resolution (Å)	37.58 – 2.04 37.58 – 2.04	Depositor EDS
% Data completeness (in resolution range)	98.3 (37.58-2.04) 98.3 (37.58-2.04)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	6.17 (at 2.03Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, R_{free}	0.191 , 0.222 0.192 , 0.221	Depositor DCC
R_{free} test set	6500 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	24.4	Xtriage
Anisotropy	0.383	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 54.9	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.006 for -h-l,k,h 0.006 for l,k,-h-l 0.021 for h,-k,-h-l 0.021 for -h-l,-k,l 0.023 for l,-k,h	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	13578	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% 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: NHE, NAG, GOL, PEG

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	0.27	0/2648	0.48	0/3596
1	C	0.26	0/2623	0.47	0/3562
1	E	0.26	0/2648	0.47	0/3595
2	B	0.26	0/1457	0.43	0/1958
2	D	0.27	0/1466	0.43	0/1969
2	F	0.26	0/1453	0.43	0/1953
All	All	0.26	0/12295	0.46	0/16633

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	2579	0	2525	17	0
1	C	2560	0	2503	21	0
1	E	2576	0	2533	23	0
2	B	1427	0	1338	11	0
2	D	1433	0	1342	11	0
2	F	1423	0	1335	10	0
3	G	28	0	25	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	28	0	25	0	0
3	I	28	0	25	0	0
4	A	13	0	17	1	0
4	C	13	0	17	0	0
4	E	13	0	17	1	0
5	A	14	0	13	0	0
5	C	28	0	26	0	0
6	A	21	0	30	0	0
6	B	14	0	20	0	0
6	C	7	0	10	0	0
6	D	14	0	20	0	0
6	E	7	0	10	0	0
7	A	6	0	8	1	0
7	C	6	0	8	1	0
7	E	12	0	16	2	0
8	A	335	0	0	3	0
8	B	124	0	0	4	0
8	C	311	0	0	2	0
8	D	151	0	0	0	0
8	E	283	0	0	2	0
8	F	124	0	0	1	0
All	All	13578	0	11863	79	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 3.

All (79) 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:314:LEU:HA	7:C:407:GOL:H32	1.68	0.76
1:A:114:ARG:NH2	8:A:501:HOH:O	2.20	0.70
1:E:294:PHE:HZ	2:F:59:MET:HG3	1.55	0.70
1:C:185:PRO:HG2	1:C:217:ILE:HG12	1.76	0.68
1:A:25:GLN:O	1:A:313:ARG:NH2	2.26	0.67
2:B:77[B]:ILE:HD11	2:D:77:ILE:HD12	1.78	0.66
1:E:244:ASN:ND2	8:E:501:HOH:O	2.23	0.65
1:E:185:PRO:HG2	1:E:217:ILE:HG12	1.79	0.64
2:B:121:LYS:NZ	8:B:302:HOH:O	2.32	0.62
1:C:294:PHE:HZ	2:D:59:MET:HG3	1.65	0.60
1:E:307:LYS:HD2	2:F:62:GLN:HG2	1.83	0.60
1:A:294:PHE:HZ	2:B:59:MET:HG3	1.68	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:216:ARG:O	1:C:220:ARG:NH2	2.35	0.59
1:C:314:LEU:HD22	2:D:100:VAL:HG21	1.86	0.58
1:A:77:ASP:HA	1:A:80:ILE:HG13	1.86	0.58
1:E:216:ARG:O	1:E:220:ARG:NH2	2.36	0.58
1:C:77:ASP:HA	1:C:80:ILE:HG13	1.86	0.57
2:B:105:GLU:OE2	8:B:301:HOH:O	2.17	0.57
2:B:50:ASN:ND2	8:B:304:HOH:O	2.40	0.55
8:C:502:HOH:O	2:D:71:ASN:ND2	2.37	0.55
1:A:307:LYS:HD2	2:B:62:GLN:HG2	1.88	0.54
1:A:244:ASN:ND2	8:A:510:HOH:O	2.39	0.54
2:F:61:THR:OG1	8:F:201:HOH:O	2.19	0.53
1:E:105:TYR:CD2	7:E:406:GOL:H11	2.44	0.53
1:E:164:ILE:O	1:E:246:GLU:HA	2.09	0.52
1:C:307:LYS:HD2	2:D:62:GLN:HG2	1.92	0.52
1:C:268:MET:HG3	1:C:284:PRO:HG3	1.92	0.52
1:E:134:GLY:HA3	1:E:153:TRP:HB3	1.92	0.51
1:C:278:ASN:ND2	8:C:510:HOH:O	2.43	0.50
1:C:164:ILE:O	1:C:246:GLU:HA	2.12	0.50
1:E:77:ASP:HA	1:E:80:ILE:HG13	1.93	0.50
1:E:284:PRO:HD3	1:E:300:LEU:O	2.12	0.49
1:E:153:TRP:CH2	4:E:401:NHE:H6'1	2.47	0.49
2:B:134:GLY:HA2	2:D:124:LEU:HD22	1.95	0.49
1:C:98:TYR:CD1	1:C:230:MET:HG2	2.48	0.49
1:C:134:GLY:HA3	1:C:153:TRP:HB3	1.94	0.48
1:A:314:LEU:HD22	2:B:100:VAL:HG21	1.94	0.48
2:D:131:LYS:NZ	2:F:127:ARG:HH11	2.10	0.48
1:A:153:TRP:CH2	4:A:401:NHE:H6'1	2.49	0.48
1:E:314:LEU:HD22	2:F:100:VAL:HG21	1.94	0.48
1:A:134:GLY:HA3	1:A:153:TRP:HB3	1.96	0.47
2:B:58:LYS:HD3	2:B:58:LYS:HA	1.70	0.47
2:B:84:MET:HB2	2:D:84[B]:MET:SD	2.54	0.47
1:A:164:ILE:O	1:A:246:GLU:HA	2.14	0.47
1:C:43:LEU:HB2	1:C:314:LEU:HB2	1.97	0.46
1:C:100:GLY:HA3	1:C:230:MET:O	2.16	0.46
1:A:268:MET:HG3	1:A:284:PRO:HG3	1.98	0.45
2:F:62:GLN:HB2	2:F:92:TRP:CD2	2.51	0.45
1:A:284:PRO:HD3	1:A:300:LEU:O	2.16	0.45
1:E:42:ILE:HA	1:E:292[A]:MET:HG2	1.98	0.45
2:F:39:GLU:H	2:F:39:GLU:CD	2.20	0.45
1:E:100:GLY:HA3	1:E:230:MET:O	2.17	0.45
1:E:98:TYR:CD1	1:E:230:MET:HG2	2.52	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:127:ARG:NH2	8:B:306:HOH:O	2.42	0.44
1:C:284:PRO:HD3	1:C:300:LEU:O	2.17	0.44
2:F:58:LYS:HD3	2:F:58:LYS:HA	1.68	0.44
2:F:39:GLU:OE1	2:F:39:GLU:N	2.48	0.43
1:A:100:GLY:HA3	1:A:230:MET:O	2.18	0.43
1:E:323:SER:HA	1:E:324:PRO:HD3	1.89	0.43
1:C:18:HIS:O	1:C:320:LEU:HD11	2.18	0.43
2:D:58:LYS:HA	2:D:58:LYS:HD3	1.67	0.43
2:D:158:ASP:OD1	2:D:160:PRO:HD2	2.18	0.43
1:C:86:TYR:CZ	1:C:282:GLN:HG2	2.54	0.42
1:E:18:HIS:O	1:E:320:LEU:HD11	2.19	0.42
1:E:96(A):LEU:HG	7:E:406:GOL:H32	2.01	0.42
2:D:83:LYS:HD3	2:D:83:LYS:HA	1.80	0.42
1:E:118:PHE:HE1	1:E:260:ILE:HG12	1.84	0.42
1:A:98:TYR:CD1	1:A:230:MET:HG2	2.54	0.42
1:C:129:SER:HB3	1:C:162:PRO:HG2	2.02	0.42
1:E:46:LYS:NZ	8:E:525:HOH:O	2.53	0.42
1:A:313:ARG:NH1	8:A:520:HOH:O	2.53	0.41
1:C:323:SER:HA	1:C:324:PRO:HD3	1.91	0.41
1:A:313:ARG:NH1	7:A:408:GOL:O2	2.53	0.41
1:C:283:THR:HG22	1:C:301:THR:HG22	2.02	0.41
1:A:203[A]:SER:OG	1:A:246:GLU:HB3	2.21	0.41
1:C:174:GLU:OE2	1:C:259:LYS:HD2	2.20	0.41
1:E:294:PHE:CZ	2:F:59:MET:HG3	2.43	0.41
1:E:203:SER:OG	1:E:246:GLU:HB3	2.22	0.40
1:E:174:GLU:OE2	1:E:259:LYS:HD2	2.21	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	324/334 (97%)	318 (98%)	6 (2%)	0	100	100
1	C	321/334 (96%)	315 (98%)	6 (2%)	0	100	100
1	E	324/334 (97%)	318 (98%)	6 (2%)	0	100	100
2	B	175/177 (99%)	172 (98%)	3 (2%)	0	100	100
2	D	176/177 (99%)	173 (98%)	3 (2%)	0	100	100
2	F	174/177 (98%)	171 (98%)	3 (2%)	0	100	100
All	All	1494/1533 (98%)	1467 (98%)	27 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	293/300 (98%)	291 (99%)	2 (1%)	84	84
1	C	290/300 (97%)	289 (100%)	1 (0%)	92	93
1	E	293/300 (98%)	290 (99%)	3 (1%)	76	75
2	B	151/151 (100%)	150 (99%)	1 (1%)	84	84
2	D	152/151 (101%)	151 (99%)	1 (1%)	84	84
2	F	151/151 (100%)	150 (99%)	1 (1%)	84	84
All	All	1330/1353 (98%)	1321 (99%)	9 (1%)	86	84

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	114	ARG
1	A	144	LYS
2	B	22	TYR
1	C	121	ILE
2	D	22	TYR
1	E	114	ARG
1	E	292[A]	MET

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type
1	E	292[B]	MET
2	F	22	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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	$\#|Z| > 2$
3	NAG	G	1	1,3	14,14,15	0.36	0	17,19,21	0.70	0
3	NAG	G	2	3	14,14,15	0.21	0	17,19,21	0.39	0
3	NAG	H	1	1,3	14,14,15	0.36	0	17,19,21	0.58	0
3	NAG	H	2	3	14,14,15	0.16	0	17,19,21	0.56	0
3	NAG	I	1	1,3	14,14,15	0.30	0	17,19,21	0.53	0
3	NAG	I	2	3	14,14,15	0.25	0	17,19,21	0.53	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	NAG	G	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	G	2	3	-	0/6/23/26	0/1/1/1
3	NAG	H	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	H	2	3	-	0/6/23/26	0/1/1/1
3	NAG	I	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	I	2	3	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	1	NAG	O5-C5-C6-O6
3	G	1	NAG	C4-C5-C6-O6
3	I	2	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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

Oligosaccharide Chain G

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain H

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain I

Bond lengths

Bond angles

Torsions

Rings
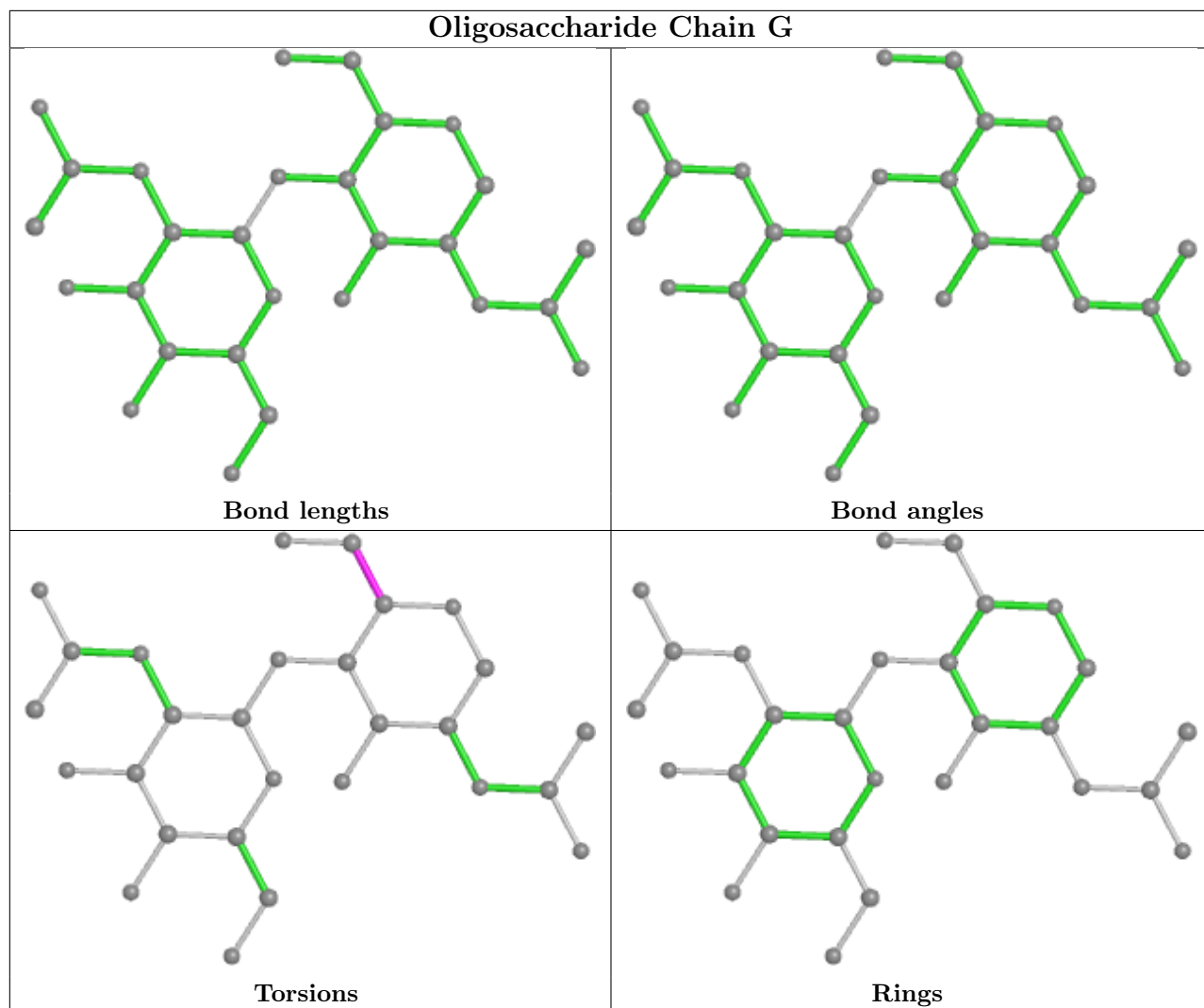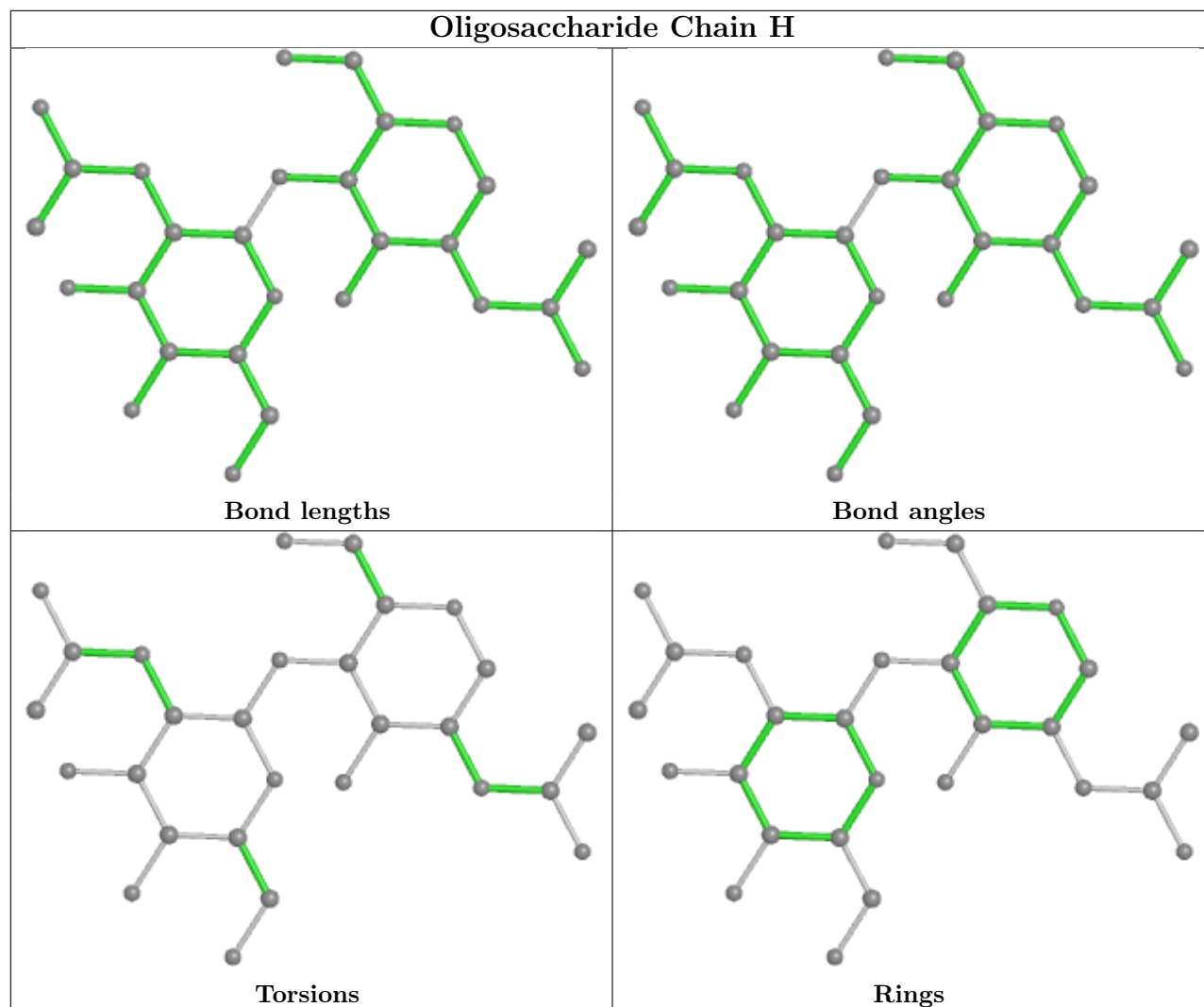
5.6 Ligand geometry ⓘ

19 ligands are modelled in this entry.

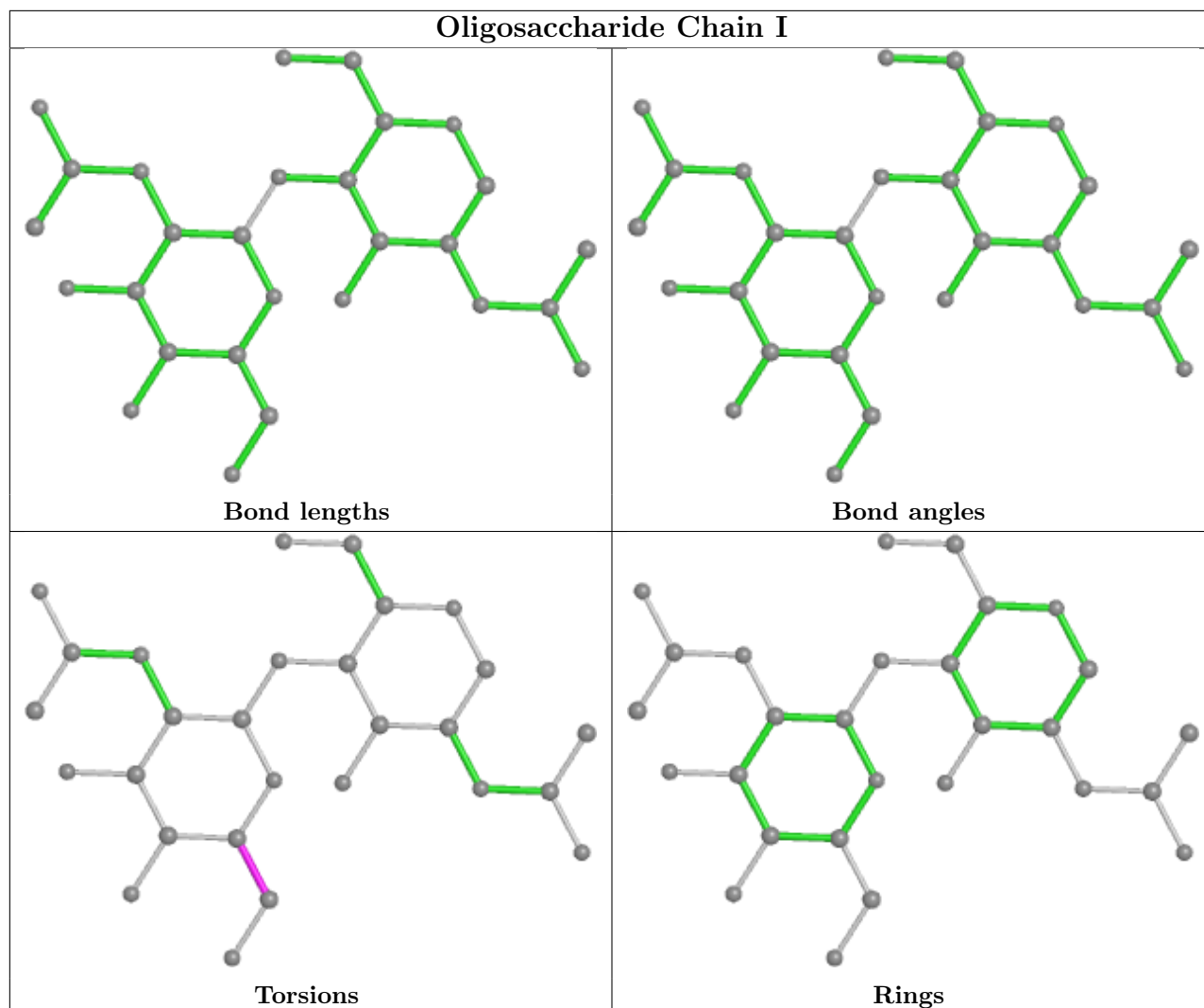In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists 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	$\#|Z| > 2$
5	NAG	C	402	1	14,14,15	0.19	0	17,19,21	0.33	0
6	PEG	A	406	-	6,6,6	0.52	0	5,5,5	0.43	0
6	PEG	C	406	-	6,6,6	0.51	0	5,5,5	0.39	0
6	PEG	E	404	-	6,6,6	0.55	0	5,5,5	0.35	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
4	NHE	C	401	-	13,13,13	1.70	3 (23%)	16,17,17	1.23	2 (12%)
7	GOL	A	408	-	5,5,5	0.92	0	5,5,5	0.96	0
7	GOL	C	407	-	5,5,5	0.69	0	5,5,5	0.75	0
6	PEG	A	405	-	6,6,6	0.34	0	5,5,5	0.47	0
7	GOL	E	405	-	5,5,5	0.87	0	5,5,5	1.00	0
4	NHE	E	401	-	13,13,13	1.74	3 (23%)	16,17,17	1.22	2 (12%)
6	PEG	A	407	-	6,6,6	0.52	0	5,5,5	0.37	0
5	NAG	C	405	1	14,14,15	0.29	0	17,19,21	0.49	0
6	PEG	D	202	-	6,6,6	0.57	0	5,5,5	0.22	0
7	GOL	E	406	-	5,5,5	0.80	0	5,5,5	1.18	1 (20%)
6	PEG	D	201	-	6,6,6	0.50	0	5,5,5	0.47	0
4	NHE	A	401	-	13,13,13	1.72	3 (23%)	16,17,17	1.18	1 (6%)
6	PEG	B	202	-	6,6,6	0.30	0	5,5,5	1.16	0
6	PEG	B	201	-	6,6,6	0.53	0	5,5,5	0.65	0
5	NAG	A	404	1	14,14,15	0.20	0	17,19,21	0.47	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
5	NAG	C	402	1	-	2/6/23/26	0/1/1/1
6	PEG	A	406	-	-	2/4/4/4	-
6	PEG	C	406	-	-	2/4/4/4	-
6	PEG	E	404	-	-	2/4/4/4	-
4	NHE	C	401	-	-	5/7/15/15	0/1/1/1
7	GOL	A	408	-	-	2/4/4/4	-
7	GOL	C	407	-	-	2/4/4/4	-
6	PEG	A	405	-	-	2/4/4/4	-
7	GOL	E	405	-	-	2/4/4/4	-
4	NHE	E	401	-	-	4/7/15/15	0/1/1/1
6	PEG	A	407	-	-	3/4/4/4	-
5	NAG	C	405	1	-	0/6/23/26	0/1/1/1
6	PEG	D	202	-	-	1/4/4/4	-
7	GOL	E	406	-	-	0/4/4/4	-
6	PEG	D	201	-	-	2/4/4/4	-
4	NHE	A	401	-	-	5/7/15/15	0/1/1/1
6	PEG	B	202	-	-	3/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	PEG	B	201	-	-	0/4/4/4	-
5	NAG	A	404	1	-	0/6/23/26	0/1/1/1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	401	NHE	C2-S	4.28	1.83	1.77
4	A	401	NHE	C2-S	4.23	1.83	1.77
4	C	401	NHE	C2-S	4.06	1.83	1.77
4	A	401	NHE	O2-S	3.01	1.53	1.45
4	E	401	NHE	O1-S	2.98	1.53	1.45
4	C	401	NHE	O1-S	2.97	1.53	1.45
4	E	401	NHE	O2-S	2.92	1.53	1.45
4	A	401	NHE	O1-S	2.82	1.53	1.45
4	C	401	NHE	O2-S	2.81	1.53	1.45

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	401	NHE	O1-S-C2	3.14	110.70	106.92
4	A	401	NHE	O2-S-C2	2.96	110.48	106.92
4	E	401	NHE	O1-S-C2	2.48	109.90	106.92
4	E	401	NHE	O3-S-C2	2.30	109.48	105.77
7	E	406	GOL	C3-C2-C1	-2.12	103.45	111.70
4	C	401	NHE	O3-S-C2	2.09	109.14	105.77

There are no chirality outliers.

All (39) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	401	NHE	C2'-C1'-N-C1
4	C	401	NHE	C2'-C1'-N-C1
4	C	401	NHE	C1-C2-S-O2
7	A	408	GOL	C1-C2-C3-O3
7	C	407	GOL	C1-C2-C3-O3
7	E	405	GOL	C1-C2-C3-O3
6	A	406	PEG	O2-C3-C4-O4
6	A	407	PEG	O2-C3-C4-O4
6	B	202	PEG	O2-C3-C4-O4
4	C	401	NHE	C1-C2-S-O3
4	E	401	NHE	C1-C2-S-O3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	A	405	PEG	O1-C1-C2-O2
4	A	401	NHE	C2-C1-N-C1'
7	A	408	GOL	O2-C2-C3-O3
7	E	405	GOL	O2-C2-C3-O3
6	B	202	PEG	O1-C1-C2-O2
6	C	406	PEG	O1-C1-C2-O2
7	C	407	GOL	O2-C2-C3-O3
5	C	402	NAG	C4-C5-C6-O6
4	C	401	NHE	C2-C1-N-C1'
4	E	401	NHE	C2-C1-N-C1'
6	A	407	PEG	C4-C3-O2-C2
6	C	406	PEG	C4-C3-O2-C2
5	C	402	NAG	O5-C5-C6-O6
6	D	202	PEG	C1-C2-O2-C3
6	E	404	PEG	O1-C1-C2-O2
4	A	401	NHE	C1-C2-S-O3
4	A	401	NHE	C1-C2-S-O1
4	A	401	NHE	C1-C2-S-O2
4	C	401	NHE	C1-C2-S-O1
4	E	401	NHE	C1-C2-S-O1
4	E	401	NHE	C1-C2-S-O2
6	E	404	PEG	C4-C3-O2-C2
6	D	201	PEG	O2-C3-C4-O4
6	B	202	PEG	C4-C3-O2-C2
6	D	201	PEG	C1-C2-O2-C3
6	A	406	PEG	O1-C1-C2-O2
6	A	407	PEG	O1-C1-C2-O2
6	A	405	PEG	O2-C3-C4-O4

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	408	GOL	1	0
7	C	407	GOL	1	0
4	E	401	NHE	1	0
7	E	406	GOL	2	0
4	A	401	NHE	1	0

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(Å²)	Q<0.9
1	A	324/334 (97%)	-0.25	5 (1%)	73	76	12, 25, 40, 70	0
1	C	323/334 (96%)	-0.20	3 (0%)	84	86	12, 28, 44, 70	0
1	E	323/334 (96%)	-0.13	4 (1%)	79	81	12, 27, 45, 77	0
2	B	176/177 (99%)	0.05	1 (0%)	89	91	11, 32, 49, 65	0
2	D	176/177 (99%)	0.01	5 (2%)	53	58	11, 30, 49, 62	0
2	F	175/177 (98%)	0.57	18 (10%)	6	6	11, 38, 59, 73	0
All	All	1497/1533 (97%)	-0.05	36 (2%)	59	63	11, 29, 50, 77	0

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	324	PRO	5.3
1	A	324	PRO	4.5
1	E	324	PRO	4.4
2	F	26	HIS	4.2
1	E	323	SER	4.1
2	F	61	THR	3.7
2	F	174	SER	3.7
2	F	18	VAL	3.6
1	C	173	GLN	3.4
2	F	31	GLY	3.3
2	F	27	SER	3.3
2	F	38	LYS	3.3
1	E	313	ARG	3.3
1	A	323	SER	3.2
2	F	29	GLU	3.2
2	F	60	ASN	3.1
2	F	147	GLU	3.1
1	A	291	SER	3.1
1	C	323	SER	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	21	ASN	2.8
2	D	29	GLU	2.8
1	A	21	ASN	2.8
2	F	33	GLY	2.7
2	B	18	VAL	2.5
2	F	148	CYS	2.3
2	F	145	ASP	2.3
2	D	31	GLY	2.3
2	F	172	GLU	2.3
2	D	147	GLU	2.1
2	F	134	GLY	2.1
2	F	154	ASN	2.1
2	F	32	SER	2.1
2	D	18	VAL	2.1
1	A	173	GLN	2.1
2	D	33	GLY	2.0
2	F	175	SER	2.0

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

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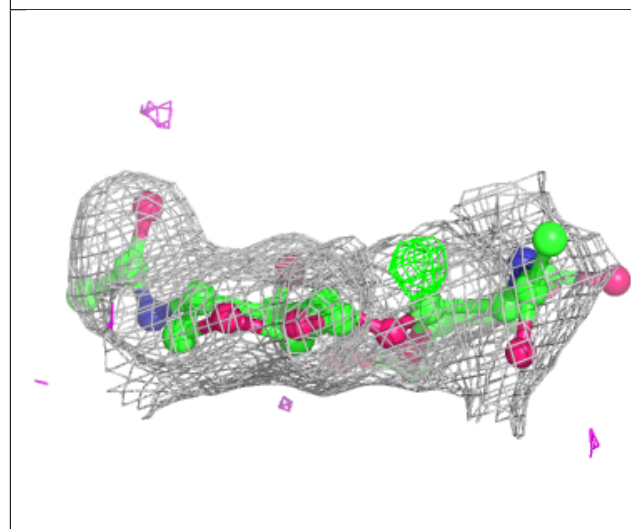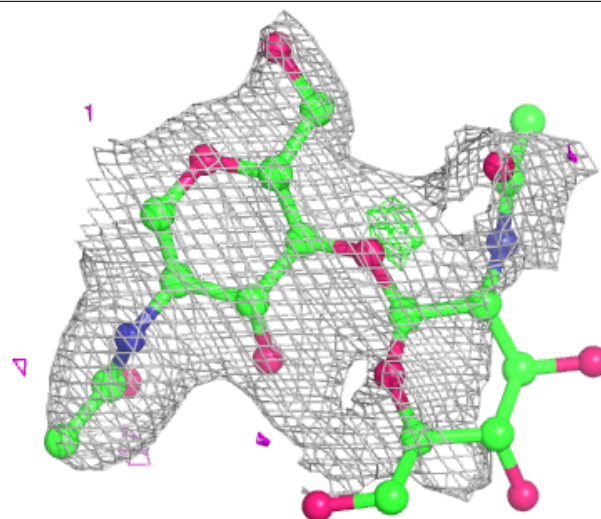
6.3 Carbohydrates ⓘ

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
3	NAG	G	2	14/15	0.68	0.45	82,90,101,104	0
3	NAG	H	2	14/15	0.77	0.37	63,76,86,87	0
3	NAG	I	2	14/15	0.83	0.30	58,71,79,80	0
3	NAG	G	1	14/15	0.92	0.15	37,48,61,70	0
3	NAG	I	1	14/15	0.92	0.09	34,41,49,57	0
3	NAG	H	1	14/15	0.92	0.12	26,34,48,55	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.
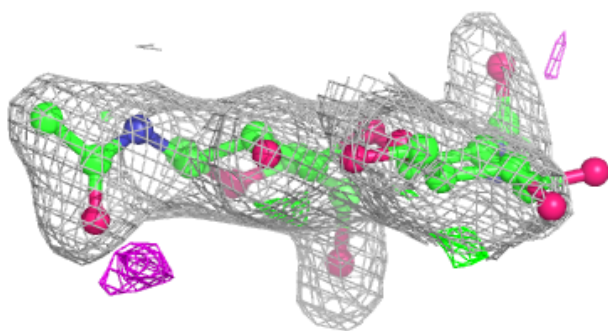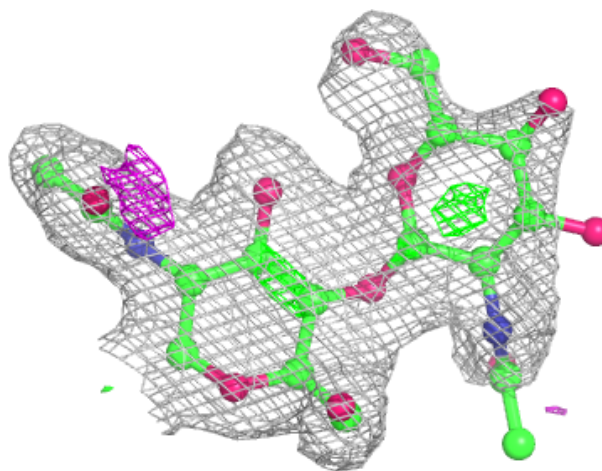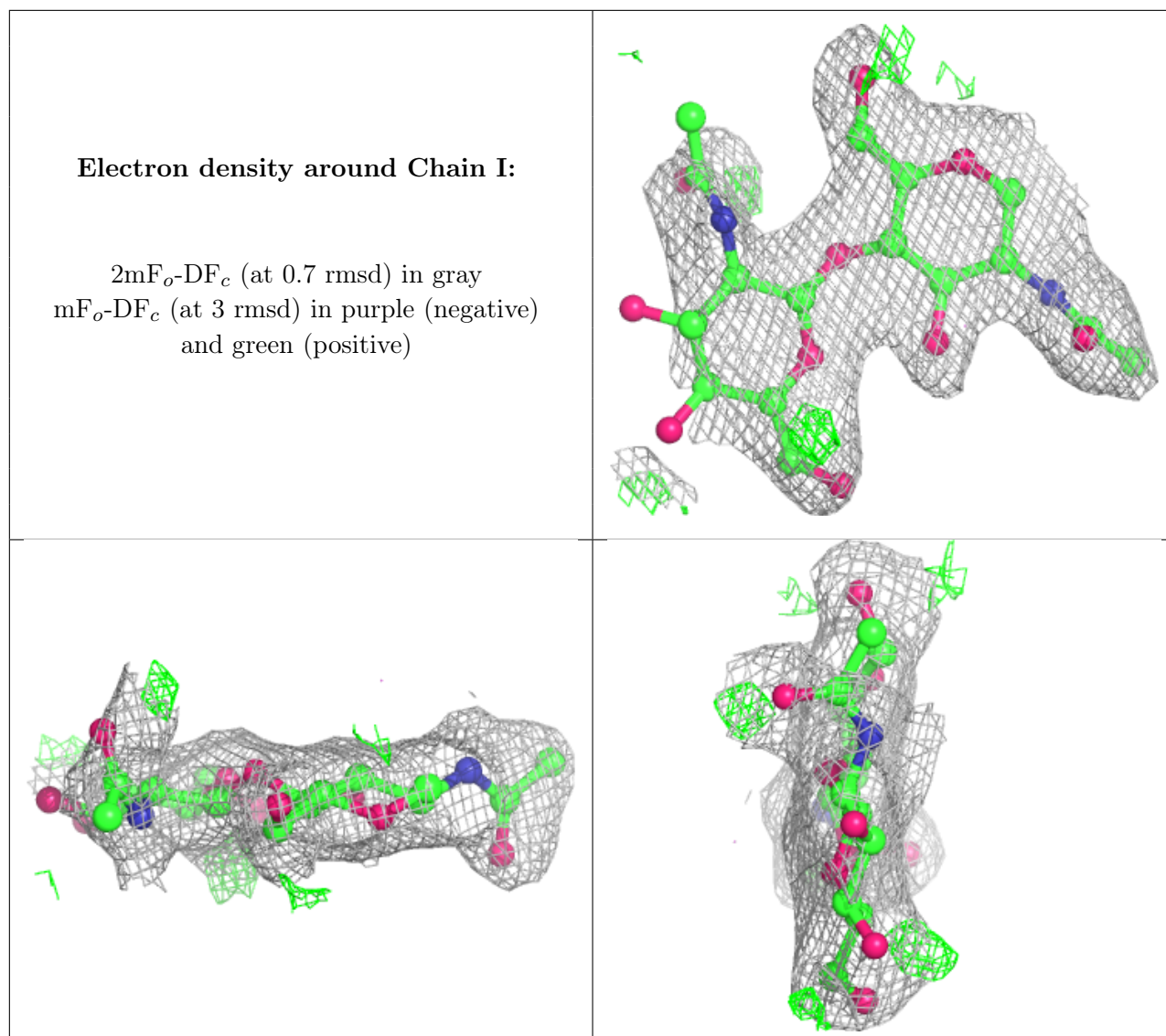
Electron density around Chain G:

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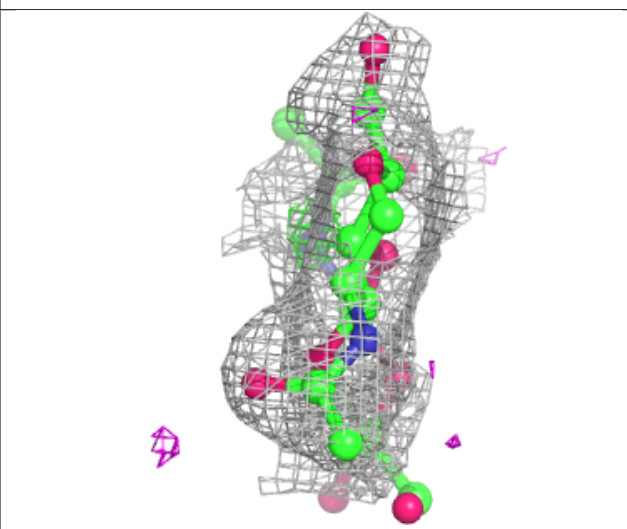
Electron density around Chain H:

$2mF_o$-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 I:

$2mF_o\text{-}DF_c$ (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
5	NAG	C	405	14/15	0.64	0.38	59,86,91,93	0
5	NAG	C	402	14/15	0.66	0.50	102,107,108,108	0
4	NHE	C	401	13/13	0.69	0.28	34,47,73,87	0
6	PEG	B	202	7/7	0.69	0.20	30,36,41,42	0
6	PEG	A	405	7/7	0.74	0.29	30,36,41,42	0
5	NAG	A	404	14/15	0.76	0.23	66,71,74,75	0
6	PEG	A	406	7/7	0.77	0.15	38,47,50,54	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	PEG	D	202	7/7	0.80	0.31	31,43,54,55	0
6	PEG	E	404	7/7	0.80	0.17	46,48,53,58	0
7	GOL	C	407	6/6	0.81	0.25	30,31,43,43	0
6	PEG	C	406	7/7	0.83	0.22	36,38,44,53	0
6	PEG	A	407	7/7	0.83	0.17	40,44,50,51	0
7	GOL	E	406	6/6	0.87	0.14	30,31,43,43	0
4	NHE	E	401	13/13	0.88	0.21	33,44,62,84	0
6	PEG	D	201	7/7	0.89	0.11	42,44,48,53	0
7	GOL	E	405	6/6	0.90	0.13	33,39,41,44	0
7	GOL	A	408	6/6	0.92	0.15	25,36,43,44	0
6	PEG	B	201	7/7	0.93	0.09	30,36,41,42	0
4	NHE	A	401	13/13	0.97	0.09	20,29,37,42	0

6.5 Other polymers ⓘ

There are no such residues in this entry.