



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 07:19 AM UTC

PDB ID : 4GXX / pdb_00004gxx
Title : Crystal structure of the "avianized" 1918 influenza virus hemagglutinin
Authors : Ekiert, D.C.; Wilson, I.A.
Deposited on : 2012-09-04
Resolution : 1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

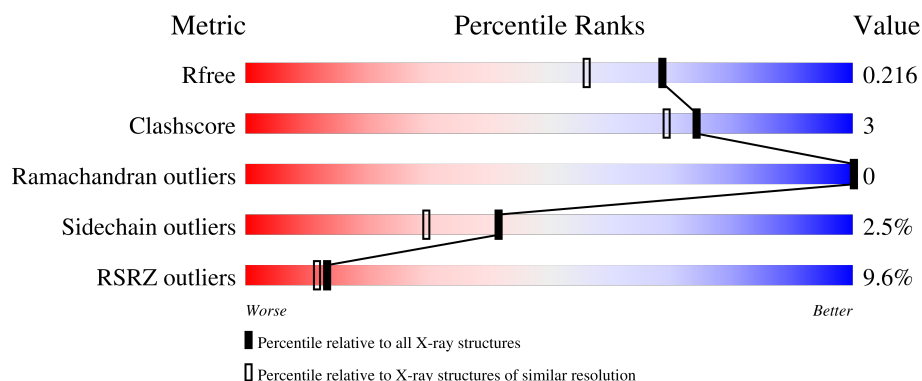
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	7662 (1.80-1.80)
Clashscore	190562	8479 (1.80-1.80)
Ramachandran outliers	187476	8391 (1.80-1.80)
Sidechain outliers	187428	8390 (1.80-1.80)
RSRZ outliers	180081	7663 (1.80-1.80)

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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	331	3% 90% 7% ..
1	C	331	5% 88% 9% .
1	E	331	5% 89% 9% .
2	B	176	22% 86% 10% ..
2	D	176	24% 85% 12% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	176	9% 90% 7%
3	G	4	100%
4	H	5	60% 40%
5	I	3	100%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12942 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 HA1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	323	Total	C	N	O	S	0	5	0
			2545	1604	440	490	11			
1	C	323	Total	C	N	O	S	0	4	0
			2536	1598	438	489	11			
1	E	325	Total	C	N	O	S	0	8	0
			2573	1622	442	498	11			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	expression tag	UNP Q9WFX3
A	8	ASP	-	expression tag	UNP Q9WFX3
A	9	PRO	-	expression tag	UNP Q9WFX3
A	10	GLY	-	expression tag	UNP Q9WFX3
A	190	GLU	ASP	engineered mutation	UNP Q9WFX3
A	225	GLY	ASP	engineered mutation	UNP Q9WFX3
C	7	ALA	-	expression tag	UNP Q9WFX3
C	8	ASP	-	expression tag	UNP Q9WFX3
C	9	PRO	-	expression tag	UNP Q9WFX3
C	10	GLY	-	expression tag	UNP Q9WFX3
C	190	GLU	ASP	engineered mutation	UNP Q9WFX3
C	225	GLY	ASP	engineered mutation	UNP Q9WFX3
E	7	ALA	-	expression tag	UNP Q9WFX3
E	8	ASP	-	expression tag	UNP Q9WFX3
E	9	PRO	-	expression tag	UNP Q9WFX3
E	10	GLY	-	expression tag	UNP Q9WFX3
E	190	GLU	ASP	engineered mutation	UNP Q9WFX3
E	225	GLY	ASP	engineered mutation	UNP Q9WFX3

- Molecule 2 is a protein called Hemagglutinin HA2 chain.

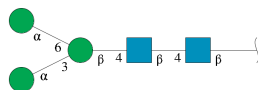
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	170	Total	C	N	O	S	0	2	0
			1387	867	241	273	6			
2	D	170	Total	C	N	O	S	0	2	0
			1385	867	239	273	6			
2	F	171	Total	C	N	O	S	0	3	0
			1399	873	241	279	6			

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



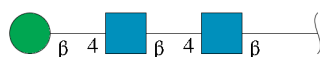
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	G	4	Total	C	N	O	0	0	0
			50	28	2	20			

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	H	5	Total	C	N	O	0	0	0
			61	34	2	25			

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



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

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	205	Total	O	0	0
			205	205		
7	B	66	Total	O	0	0
			66	66		
7	C	222	Total	O	0	0
			222	222		
7	D	71	Total	O	0	0
			71	71		
7	E	233	Total	O	0	0
			233	233		

Continued on next page...

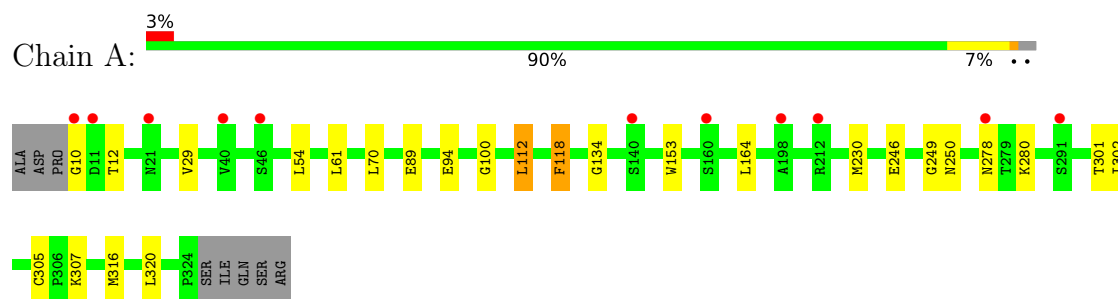
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	F	100	Total 100	O 100	0	0

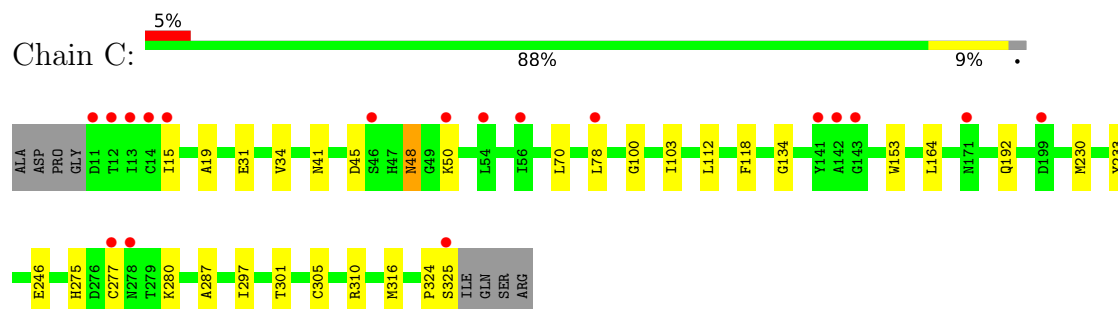
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

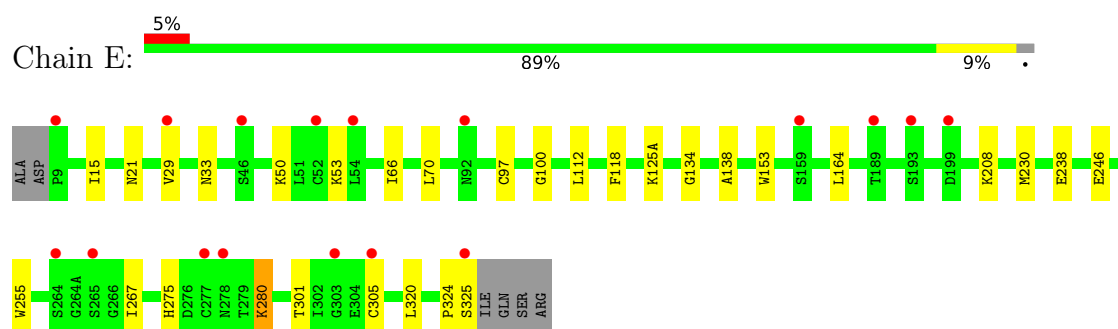
- Molecule 1: Hemagglutinin HA1 chain



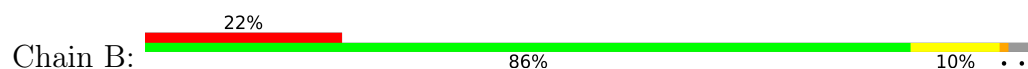
- Molecule 1: Hemagglutinin HA1 chain

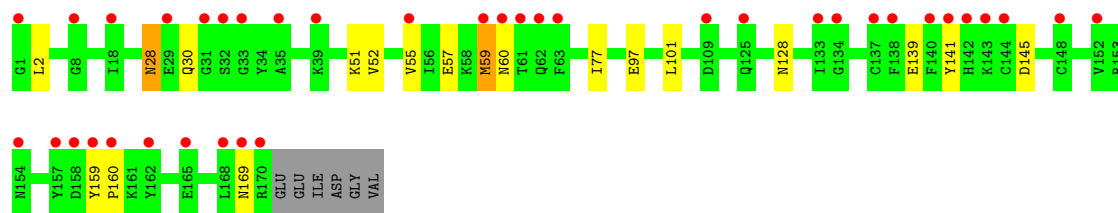


- Molecule 1: Hemagglutinin HA1 chain

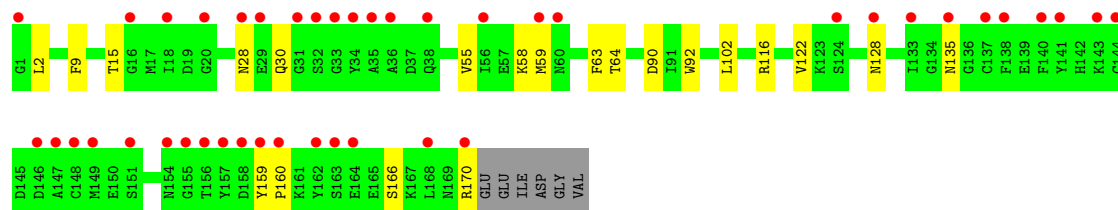
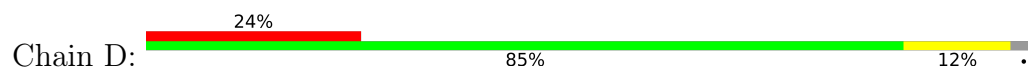


- Molecule 2: Hemagglutinin HA2 chain





● Molecule 2: Hemagglutinin HA2 chain



● Molecule 2: Hemagglutinin HA2 chain



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



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



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



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	71.79Å 241.51Å 72.03Å 90.00° 119.77° 90.00°	Depositor
Resolution (Å)	35.90 – 1.80 35.90 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.4 (35.90-1.80) 96.4 (35.90-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.47 (at 1.79Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.180 , 0.209 0.189 , 0.216	Depositor DCC
R_{free} test set	9616 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å ²)	29.7	Xtriage
Anisotropy	0.238	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 52.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.005 for -h-l,k,h 0.005 for l,k,-h-l 0.025 for h,-k,-h-l 0.023 for -h-l,-k,l 0.023 for l,-k,h	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	12942	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, NAG, MAN

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	0.69	0/2609	0.81	1/3551 (0.0%)
1	C	0.71	0/2600	0.81	0/3540
1	E	0.74	0/2638	0.83	0/3594
2	B	0.55	0/1414	0.74	0/1903
2	D	0.53	0/1412	0.76	0/1900
2	F	0.63	0/1426	0.83	0/1920
All	All	0.67	0/12099	0.80	1/16408 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	12	THR	N-CA-C	5.66	118.01	109.23

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	2545	0	2468	14	0
1	C	2536	0	2458	20	0
1	E	2573	0	2494	21	0
2	B	1387	0	1310	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1385	0	1310	12	0
2	F	1399	0	1312	8	0
3	G	50	0	43	0	0
4	H	61	0	52	0	0
5	I	39	0	34	0	0
6	A	14	0	13	0	0
6	C	28	0	26	0	0
6	E	14	0	13	1	0
6	F	14	0	13	0	0
7	A	205	0	0	0	0
7	B	66	0	0	0	0
7	C	222	0	0	4	0
7	D	71	0	0	2	0
7	E	233	0	0	2	0
7	F	100	0	0	1	0
All	All	12942	0	11546	74	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 (74) 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:E:53:LYS:NZ	7:E:707:HOH:O	2.22	0.67
1:C:192:GLN:HG2	7:C:719:HOH:O	1.97	0.64
1:E:50:LYS:HD2	1:E:275:HIS:CG	2.33	0.64
1:E:21:ASN:HB2	6:E:401:NAG:H82	1.80	0.63
1:C:70:LEU:HD11	1:C:112:LEU:HD11	1.79	0.63
1:A:301:THR:HB	1:A:305:CYS:SG	2.39	0.62
1:E:66[B]:ILE:CD1	1:E:267:ILE:HD13	2.31	0.61
1:C:50:LYS:HD2	1:C:275:HIS:CG	2.35	0.61
1:C:41:ASN:ND2	7:C:544:HOH:O	2.33	0.60
1:E:29:VAL:CG2	2:F:102:LEU:HD23	2.32	0.59
1:E:70:LEU:HD11	1:E:112:LEU:HD11	1.85	0.59
1:E:324:PRO:O	1:E:325:SER:C	2.45	0.58
1:C:310:ARG:NH1	2:D:90:ASP:OD1	2.34	0.58
1:A:10:GLY:HA3	2:B:139:GLU:OE1	2.04	0.57
1:A:70:LEU:HD11	1:A:112:LEU:HD13	1.87	0.56
7:D:269:HOH:O	2:F:59:MET:SD	2.58	0.56
2:D:9:PHE:O	2:D:135:ASN:HA	2.07	0.55
2:F:30:GLN:HE22	2:F:146:ASP:H	1.54	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:134:GLY:HA3	1:A:153:TRP:HB3	1.90	0.53
1:C:277:CYS:N	7:C:704:HOH:O	2.40	0.53
2:D:59:MET:HE1	2:D:92:TRP:CZ3	2.44	0.53
1:E:280:LYS:HB2	7:E:730:HOH:O	2.09	0.51
1:E:15:ILE:HD12	1:E:15:ILE:N	2.26	0.50
2:D:159:TYR:HB3	2:D:160:PRO:HD3	1.93	0.50
2:B:159:TYR:HB3	2:B:160:PRO:HD3	1.92	0.50
1:C:100:GLY:HA3	1:C:230:MET:O	2.11	0.50
1:A:100:GLY:HA3	1:A:230:MET:O	2.12	0.50
1:C:301:THR:HB	1:C:305:CYS:SG	2.51	0.50
1:E:134:GLY:HA3	1:E:153:TRP:HB3	1.95	0.49
1:A:316:MET:SD	2:B:52:VAL:HG22	2.52	0.49
2:B:51:LYS:HG3	1:E:29:VAL:CG1	2.43	0.48
1:E:301:THR:HB	1:E:305:CYS:SG	2.53	0.48
1:E:66[B]:ILE:HD11	1:E:267:ILE:HD13	1.96	0.48
2:D:59:MET:HE3	2:D:59:MET:HA	1.96	0.48
1:C:324:PRO:O	1:C:325:SER:C	2.56	0.47
2:D:116:ARG:HD3	7:D:261:HOH:O	2.15	0.47
1:A:316:MET:HE3	1:A:316:MET:HB2	1.78	0.47
1:C:45:ASP:C	1:C:297[A]:ILE:HD11	2.39	0.46
1:C:134:GLY:HA3	1:C:153:TRP:HB3	1.96	0.46
1:E:125(A):LYS:HD2	1:E:255:TRP:CH2	2.51	0.46
2:B:97:GLU:HB3	2:D:58:LYS:HE3	1.99	0.45
1:E:164:LEU:O	1:E:246:GLU:HA	2.16	0.45
1:C:164:LEU:O	1:C:246:GLU:HA	2.16	0.45
2:B:55:VAL:O	2:B:59:MET:HG2	2.17	0.45
1:E:50:LYS:HD2	1:E:275:HIS:CD2	2.52	0.44
1:A:249:GLY:C	1:A:250:ASN:HD22	2.25	0.44
1:E:208:LYS:HD2	1:E:238:GLU:OE2	2.18	0.44
1:A:54:LEU:HD11	1:A:302:ILE:CG2	2.48	0.44
1:C:15:ILE:HD11	2:D:122:VAL:HG21	1.99	0.43
2:F:167:LYS:NZ	7:F:390:HOH:O	2.48	0.43
2:B:141:TYR:O	2:B:169:ASN:ND2	2.50	0.43
1:E:100:GLY:HA3	1:E:230:MET:O	2.19	0.43
1:A:70:LEU:HD11	1:A:112:LEU:CD1	2.48	0.42
1:C:48:ASN:ND2	1:C:287:ALA:H	2.17	0.42
1:A:61:LEU:HD12	1:A:89:GLU:HG2	2.00	0.42
2:B:30:GLN:NE2	2:B:145:ASP:HB2	2.34	0.42
2:D:63:PHE:O	2:D:64:THR:C	2.62	0.42
1:E:97:CYS:HB2	1:E:138:ALA:O	2.19	0.42
1:A:118:PHE:C	1:A:118:PHE:CD1	2.98	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:164:LEU:O	1:A:246:GLU:HA	2.20	0.42
2:B:77[A]:ILE:HD13	2:F:77[A]:ILE:HD11	2.02	0.42
1:E:15:ILE:N	1:E:15:ILE:CD1	2.81	0.42
1:C:78:LEU:HG	7:C:722:HOH:O	2.19	0.41
1:C:31:GLU:HG3	1:C:34:VAL:CG2	2.50	0.41
2:B:28:ASN:HD22	2:B:28:ASN:C	2.28	0.41
1:C:48:ASN:C	1:C:48:ASN:HD22	2.28	0.41
2:F:30:GLN:NE2	2:F:145:ASP:HB2	2.34	0.41
1:C:19:ALA:O	2:D:15:THR:HA	2.21	0.41
1:C:103:ILE:HG12	1:C:233:TYR:CE2	2.55	0.41
2:D:166:SER:O	2:D:170:ARG:HB2	2.20	0.41
1:A:54:LEU:HD11	1:A:302:ILE:HG22	2.01	0.41
1:E:29:VAL:HG21	2:F:102:LEU:HD23	2.03	0.41
2:F:59:MET:HE3	2:F:59:MET:HA	2.03	0.40
1:C:316:MET:HE1	2:D:55:VAL:HG21	2.03	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	326/331 (98%)	319 (98%)	7 (2%)	0	100	100
1	C	325/331 (98%)	317 (98%)	8 (2%)	0	100	100
1	E	331/331 (100%)	321 (97%)	10 (3%)	0	100	100
2	B	170/176 (97%)	168 (99%)	2 (1%)	0	100	100
2	D	170/176 (97%)	168 (99%)	2 (1%)	0	100	100
2	F	172/176 (98%)	169 (98%)	3 (2%)	0	100	100
All	All	1494/1521 (98%)	1462 (98%)	32 (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	282/284 (99%)	274 (97%)	8 (3%)	38	26
1	C	282/284 (99%)	279 (99%)	3 (1%)	65	60
1	E	287/284 (101%)	283 (99%)	4 (1%)	59	52
2	B	147/150 (98%)	140 (95%)	7 (5%)	23	11
2	D	147/150 (98%)	142 (97%)	5 (3%)	32	20
2	F	149/150 (99%)	143 (96%)	6 (4%)	28	15
All	All	1294/1302 (99%)	1261 (97%)	33 (3%)	42	28

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	VAL
1	A	94	GLU
1	A	112	LEU
1	A	118	PHE
1	A	278	ASN
1	A	280	LYS
1	A	307	LYS
1	A	320	LEU
2	B	2	LEU
2	B	28	ASN
2	B	57	GLU
2	B	59	MET
2	B	60	ASN
2	B	101	LEU
2	B	128	ASN
1	C	48	ASN
1	C	118	PHE
1	C	280	LYS
2	D	2	LEU
2	D	28	ASN
2	D	30	GLN
2	D	102	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	128	ASN
1	E	33	ASN
1	E	118	PHE
1	E	280	LYS
1	E	320	LEU
2	F	2	LEU
2	F	32	SER
2	F	72[A]	ASN
2	F	72[B]	ASN
2	F	128	ASN
2	F	143	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (30) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	191	GLN
1	A	250	ASN
1	A	275	HIS
2	B	25	HIS
2	B	28	ASN
2	B	50	ASN
2	B	125	GLN
2	B	128	ASN
2	B	129	ASN
1	C	33	ASN
1	C	48	ASN
1	C	92	ASN
1	C	150	ASN
1	C	191	GLN
1	C	196	GLN
1	C	250	ASN
2	D	28	ASN
2	D	129	ASN
2	D	142	HIS
1	E	150	ASN
1	E	191	GLN
1	E	196	GLN
1	E	197	ASN
1	E	250	ASN
2	F	25	HIS
2	F	30	GLN
2	F	43	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	125	GLN
2	F	128	ASN
2	F	129	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 ⓘ

12 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	3,1	14,14,15	0.80	1 (7%)	17,19,21	1.25	2 (11%)
3	NAG	G	2	3	14,14,15	0.68	1 (7%)	17,19,21	1.51	2 (11%)
3	BMA	G	3	3	11,11,12	1.24	1 (9%)	15,15,17	0.87	0
3	MAN	G	4	3	11,11,12	0.46	0	15,15,17	0.89	1 (6%)
4	NAG	H	1	4,1	14,14,15	0.61	0	17,19,21	1.29	1 (5%)
4	NAG	H	2	4	14,14,15	0.66	0	17,19,21	0.94	0
4	BMA	H	3	4	11,11,12	1.59	2 (18%)	15,15,17	1.20	1 (6%)
4	MAN	H	4	4	11,11,12	0.56	0	15,15,17	1.04	0
4	MAN	H	5	4	11,11,12	0.47	0	15,15,17	0.68	0
5	NAG	I	1	5,1	14,14,15	0.78	1 (7%)	17,19,21	1.04	2 (11%)
5	NAG	I	2	5	14,14,15	0.75	0	17,19,21	1.04	1 (5%)
5	BMA	I	3	5	11,11,12	1.31	2 (18%)	15,15,17	0.98	1 (6%)

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	3,1	-	0/6/23/26	0/1/1/1
3	NAG	G	2	3	-	2/6/23/26	0/1/1/1
3	BMA	G	3	3	-	0/2/19/22	0/1/1/1
3	MAN	G	4	3	-	0/2/19/22	0/1/1/1
4	NAG	H	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	H	2	4	-	0/6/23/26	0/1/1/1
4	BMA	H	3	4	-	2/2/19/22	0/1/1/1
4	MAN	H	4	4	-	0/2/19/22	0/1/1/1
4	MAN	H	5	4	-	2/2/19/22	0/1/1/1
5	NAG	I	1	5,1	-	0/6/23/26	0/1/1/1
5	NAG	I	2	5	-	1/6/23/26	0/1/1/1
5	BMA	I	3	5	-	0/2/19/22	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	3	BMA	O5-C1	-3.20	1.38	1.43
3	G	3	BMA	C2-C3	2.85	1.56	1.52
3	G	1	NAG	C1-C2	2.50	1.55	1.52
5	I	3	BMA	O5-C5	2.50	1.48	1.43
4	H	3	BMA	C2-C3	2.46	1.56	1.52
5	I	3	BMA	C4-C5	2.14	1.57	1.53
5	I	1	NAG	O5-C1	-2.05	1.40	1.43
3	G	2	NAG	O5-C1	-2.04	1.40	1.43

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	1	NAG	O5-C1-C2	-3.73	105.52	111.29
3	G	2	NAG	C1-O5-C5	3.61	117.03	112.19
3	G	1	NAG	O5-C1-C2	-3.56	105.78	111.29
4	H	3	BMA	C1-O5-C5	2.39	115.39	112.19
3	G	1	NAG	C1-C2-N2	2.38	114.18	110.43
5	I	3	BMA	C1-O5-C5	2.29	115.25	112.19
3	G	2	NAG	O7-C7-C8	-2.24	118.06	122.05
5	I	1	NAG	O5-C1-C2	-2.22	107.86	111.29
5	I	2	NAG	O5-C1-C2	-2.21	107.86	111.29
3	G	4	MAN	O5-C5-C6	2.14	111.82	107.66

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	I	1	NAG	C1-C2-N2	2.07	113.69	110.43

There are no chirality outliers.

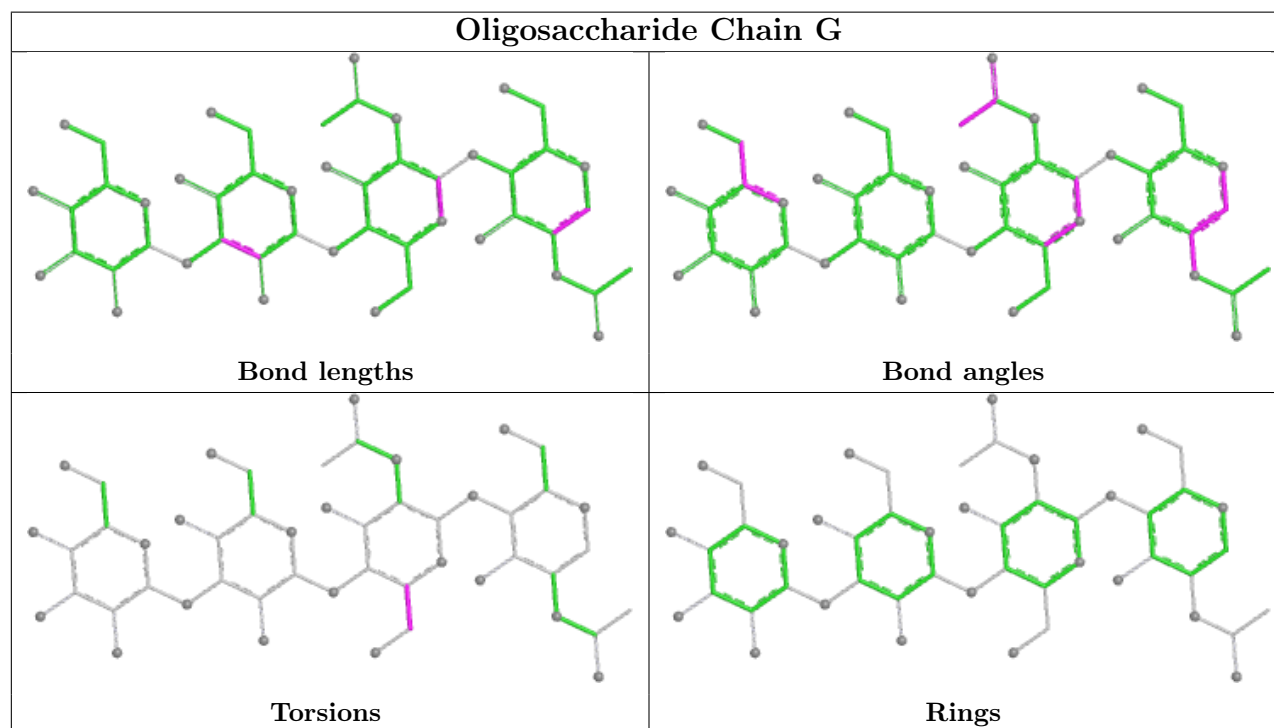
All (7) torsion outliers are listed below:

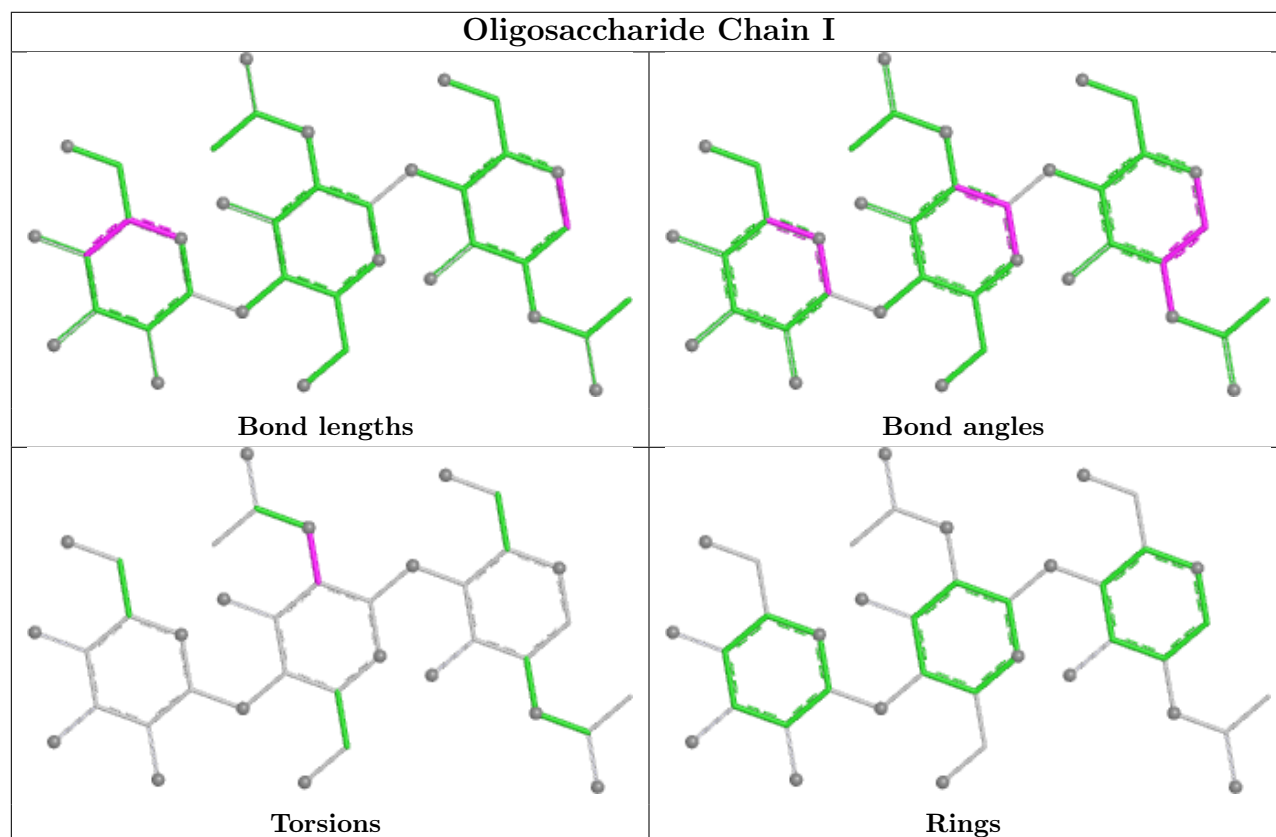
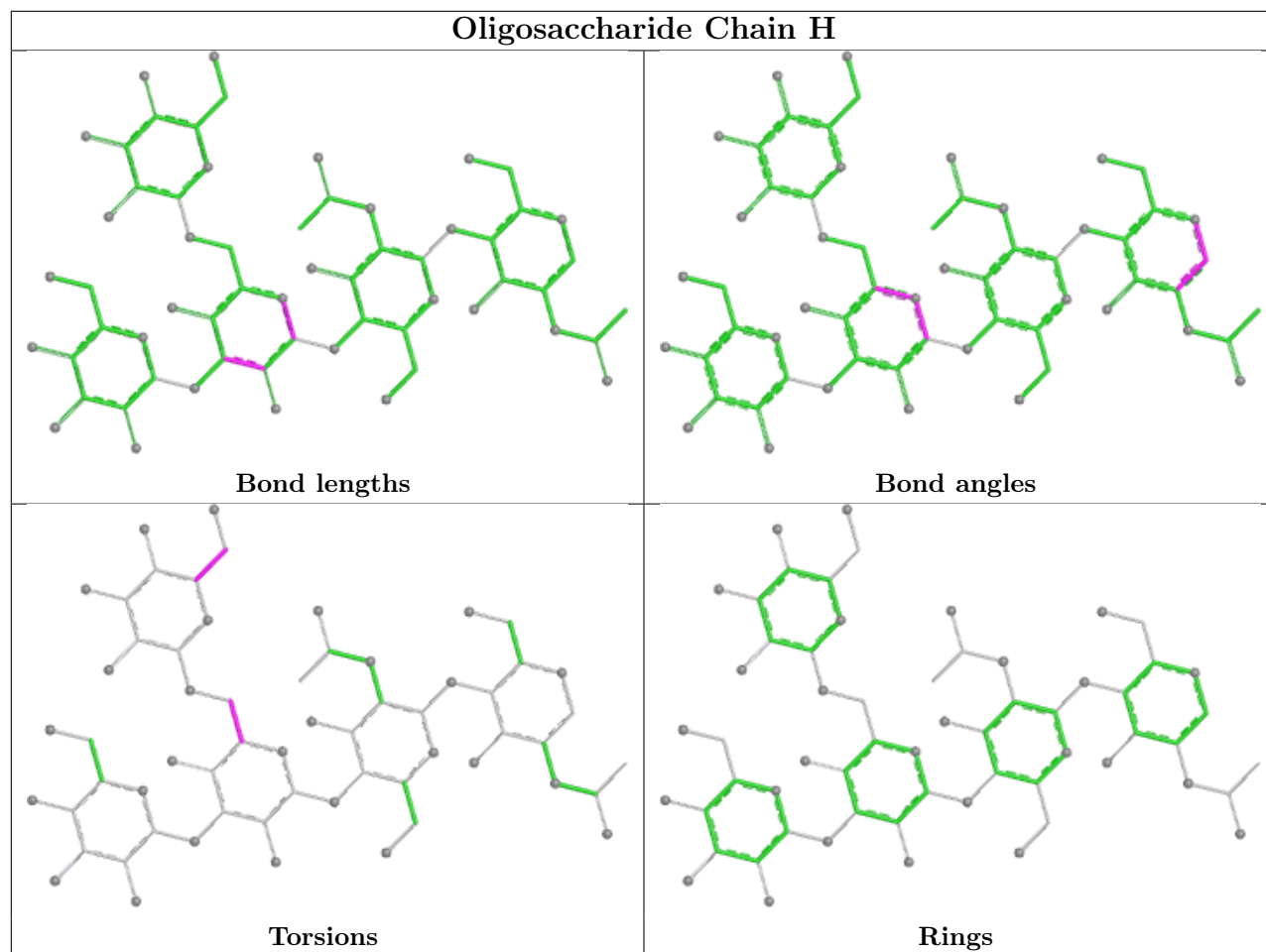
Mol	Chain	Res	Type	Atoms
4	H	3	BMA	O5-C5-C6-O6
3	G	2	NAG	O5-C5-C6-O6
4	H	3	BMA	C4-C5-C6-O6
3	G	2	NAG	C4-C5-C6-O6
4	H	5	MAN	C4-C5-C6-O6
4	H	5	MAN	O5-C5-C6-O6
5	I	2	NAG	C1-C2-N2-C7

There are no ring outliers.

No monomer is involved in short contacts.

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





5.6 Ligand geometry

5 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
6	NAG	E	401	1	14,14,15	0.52	0	17,19,21	0.97	1 (5%)
6	NAG	F	201	2	14,14,15	0.68	1 (7%)	17,19,21	1.30	3 (17%)
6	NAG	C	401	1	14,14,15	0.65	0	17,19,21	1.26	2 (11%)
6	NAG	C	407	1	14,14,15	0.51	0	17,19,21	0.80	1 (5%)
6	NAG	A	401	1	14,14,15	0.56	0	17,19,21	1.04	1 (5%)

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
6	NAG	E	401	1	-	4/6/23/26	0/1/1/1
6	NAG	F	201	2	-	2/6/23/26	0/1/1/1
6	NAG	C	401	1	-	1/6/23/26	0/1/1/1
6	NAG	C	407	1	-	2/6/23/26	0/1/1/1
6	NAG	A	401	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	F	201	NAG	C1-C2	2.02	1.55	1.52

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	401	NAG	C1-O5-C5	3.62	117.04	112.19
6	C	401	NAG	C1-O5-C5	3.13	116.38	112.19
6	F	201	NAG	C1-O5-C5	3.02	116.24	112.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	E	401	NAG	C1-O5-C5	2.80	115.94	112.19
6	C	401	NAG	O5-C1-C2	2.74	115.53	111.29
6	F	201	NAG	C3-C4-C5	-2.54	105.62	110.23
6	F	201	NAG	O5-C5-C6	2.29	112.13	107.66
6	C	407	NAG	C1-O5-C5	2.01	114.88	112.19

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	F	201	NAG	O7-C7-N2-C2
6	E	401	NAG	C8-C7-N2-C2
6	E	401	NAG	O7-C7-N2-C2
6	F	201	NAG	C8-C7-N2-C2
6	C	407	NAG	O5-C5-C6-O6
6	C	407	NAG	C4-C5-C6-O6
6	E	401	NAG	O5-C5-C6-O6
6	C	401	NAG	O5-C5-C6-O6
6	E	401	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	E	401	NAG	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	323/331 (97%)	0.54	11 (3%)	48	48	15, 44, 71, 97	5 (1%)
1	C	323/331 (97%)	0.45	18 (5%)	30	29	17, 43, 73, 103	4 (1%)
1	E	325/331 (98%)	0.39	17 (5%)	33	31	13, 41, 65, 90	8 (2%)
2	B	170/176 (96%)	1.29	38 (22%)	2	2	15, 70, 105, 115	2 (1%)
2	D	170/176 (96%)	1.31	43 (25%)	1	1	15, 69, 116, 129	2 (1%)
2	F	171/176 (97%)	0.68	15 (8%)	15	13	14, 47, 70, 109	3 (1%)
All	All	1482/1521 (97%)	0.68	142 (9%)	13	12	13, 46, 94, 129	24 (1%)

All (142) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	277	CYS	5.7
2	D	155	GLY	5.7
2	D	156	THR	4.5
2	D	148	CYS	4.5
2	D	168	LEU	4.5
2	B	133	ILE	4.4
2	D	60	ASN	4.4
2	D	144	CYS	4.3
1	E	9	PRO	4.3
1	E	278	ASN	4.2
2	D	154	ASN	4.1
2	F	61	THR	4.1
1	A	212[A]	ARG	4.0
2	B	33	GLY	4.0
1	A	10	GLY	3.9
2	B	31	GLY	3.9
2	B	63	PHE	3.7
2	B	168	LEU	3.7
1	C	142	ALA	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	151	SER	3.6
1	C	78	LEU	3.4
1	A	11	ASP	3.4
1	C	278	ASN	3.4
2	B	134	GLY	3.3
2	F	60	ASN	3.3
2	F	65	ALA	3.3
2	B	60	ASN	3.3
2	B	61	THR	3.2
2	D	147	ALA	3.2
2	B	137	CYS	3.1
1	C	12	THR	3.1
2	D	128	ASN	3.1
2	B	170	ARG	3.1
2	B	157	TYR	3.1
2	B	55	VAL	3.0
2	B	160	PRO	3.0
2	D	32	SER	3.0
2	D	140	PHE	3.0
1	C	199	ASP	3.0
1	E	325	SER	3.0
2	B	144	CYS	3.0
1	C	14	CYS	2.9
2	F	134	GLY	2.9
2	D	143	LYS	2.9
2	D	31	GLY	2.9
1	E	193[A]	SER	2.9
2	F	40[A]	SER	2.9
2	F	63	PHE	2.9
1	E	189	THR	2.9
1	E	303	GLY	2.9
1	E	199	ASP	2.9
2	B	39	LYS	2.9
2	F	62	GLN	2.8
2	F	133	ILE	2.8
2	F	66	VAL	2.8
2	B	141	TYR	2.8
2	D	133	ILE	2.8
2	B	138	PHE	2.8
2	B	152	VAL	2.8
2	D	164	GLU	2.8
2	D	137	CYS	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	46	SER	2.8
2	D	1	GLY	2.8
1	C	171[A]	ASN	2.7
1	E	92[A]	ASN	2.7
2	D	160	PRO	2.7
2	D	124	SER	2.7
2	B	18	ILE	2.7
1	A	291	SER	2.6
2	D	59	MET	2.6
1	E	29	VAL	2.6
1	E	159	SER	2.6
2	D	149	MET	2.6
1	C	143	GLY	2.5
1	E	305	CYS	2.5
1	E	46	SER	2.5
2	D	163	SER	2.5
2	D	38	GLN	2.5
1	A	21	ASN	2.5
2	F	171	GLU	2.5
1	E	54[A]	LEU	2.5
2	D	141	TYR	2.5
2	B	32	SER	2.5
1	A	40	VAL	2.5
2	D	170	ARG	2.5
1	A	46	SER	2.5
1	C	13	ILE	2.4
1	C	11	ASP	2.4
2	B	59	MET	2.4
1	C	54	LEU	2.4
2	D	138	PHE	2.4
1	C	141	TYR	2.4
1	C	15	ILE	2.4
2	B	158	ASP	2.4
2	D	158	ASP	2.4
2	B	143	LYS	2.4
2	D	18	ILE	2.4
1	A	198	ALA	2.3
2	D	159	TYR	2.3
2	B	8	GLY	2.3
2	D	29	GLU	2.3
2	F	55	VAL	2.3
2	B	142	HIS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	140	SER	2.3
2	F	72[A]	ASN	2.2
1	A	160	SER	2.2
2	F	67	GLY	2.2
2	B	169	ASN	2.2
2	D	35	ALA	2.2
2	D	146	ASP	2.2
2	B	162	TYR	2.2
2	D	28	ASN	2.2
2	D	135	ASN	2.2
2	D	20	GLY	2.2
2	D	33	GLY	2.2
2	D	56	ILE	2.2
2	B	35	ALA	2.1
1	E	277	CYS	2.1
1	C	325	SER	2.1
2	B	140	PHE	2.1
2	B	62	GLN	2.1
2	B	159	TYR	2.1
2	D	157	TYR	2.1
1	E	52	CYS	2.1
2	B	154	ASN	2.1
2	D	34	TYR	2.1
2	D	162	TYR	2.1
2	D	16	GLY	2.1
2	D	36	ALA	2.1
1	A	278	ASN	2.1
2	B	165	GLU	2.1
1	E	264	SER	2.1
1	E	265	SER	2.1
1	C	50	LYS	2.1
2	B	1	GLY	2.0
2	B	125	GLN	2.0
2	B	148	CYS	2.0
2	B	29	GLU	2.0
2	F	59	MET	2.0
1	C	56	ILE	2.0
2	B	109	ASP	2.0
2	F	39	LYS	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 ⓘ

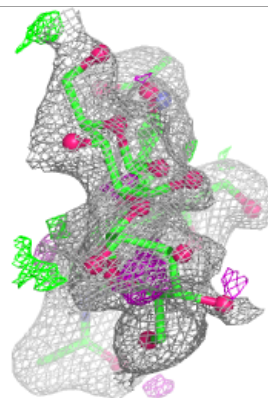
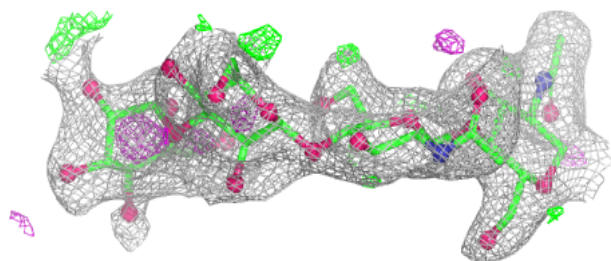
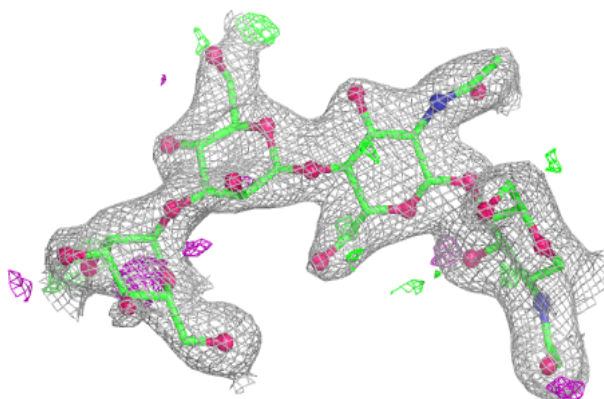
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	MAN	H	5	11/12	0.40	0.19	106,111,114,114	0
5	BMA	I	3	11/12	0.57	0.18	72,79,85,85	0
3	BMA	G	3	11/12	0.70	0.14	78,86,89,89	0
4	BMA	H	3	11/12	0.80	0.14	67,80,103,110	0
3	MAN	G	4	11/12	0.81	0.14	47,79,84,85	0
3	NAG	G	2	14/15	0.82	0.16	51,64,72,76	0
5	NAG	I	2	14/15	0.85	0.15	50,55,66,72	0
4	MAN	H	4	11/12	0.86	0.12	50,65,70,70	0
4	NAG	H	2	14/15	0.88	0.13	47,61,67,72	0
3	NAG	G	1	14/15	0.94	0.09	29,37,47,51	0
4	NAG	H	1	14/15	0.95	0.09	27,37,48,51	0
5	NAG	I	1	14/15	0.96	0.07	28,32,45,51	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.

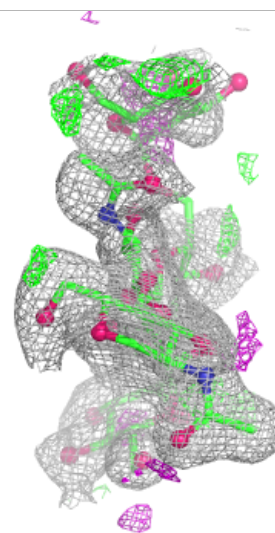
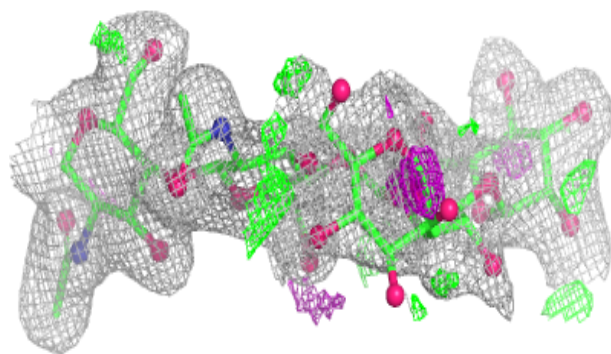
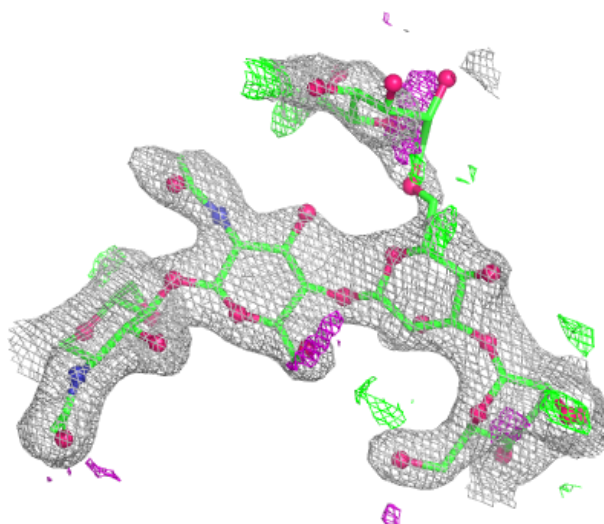
Electron density around Chain G:

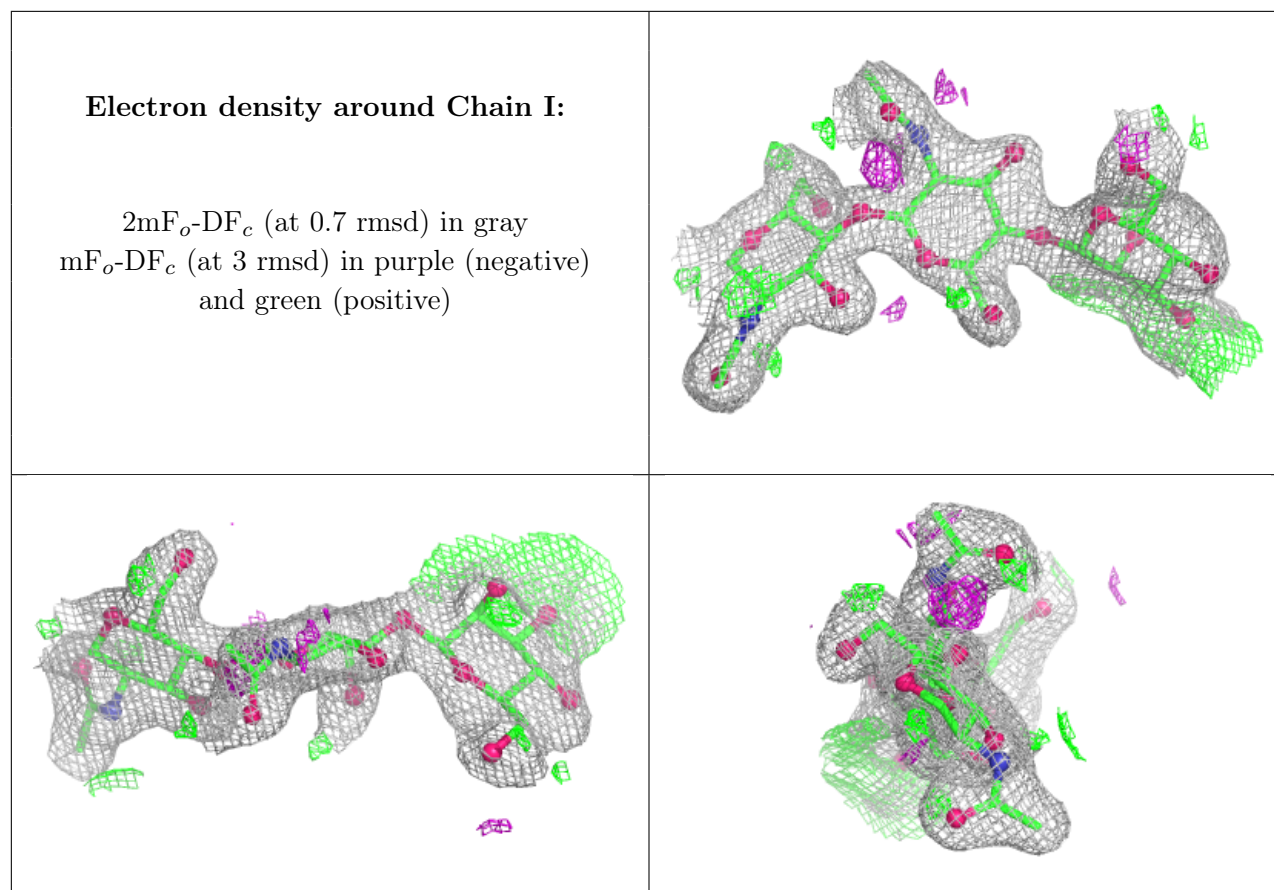
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



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)





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($\text{\AA}^2$)	Q<0.9
6	NAG	A	401	14/15	0.33	0.20	81,98,103,104	0
6	NAG	C	401	14/15	0.54	0.18	84,100,104,105	0
6	NAG	C	407	14/15	0.62	0.19	81,97,99,101	0
6	NAG	E	401	14/15	0.63	0.15	87,104,108,110	0
6	NAG	F	201	14/15	0.72	0.17	60,79,89,90	0

6.5 Other polymers ⓘ

There are no such residues in this entry.