Full wwPDB X-ray Structure Validation Report (i)

Sep 7, 2023 – 05:55 AM EDT

PDB ID	:	4GXX
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 (i) 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 (i)) were used in the production of this report:

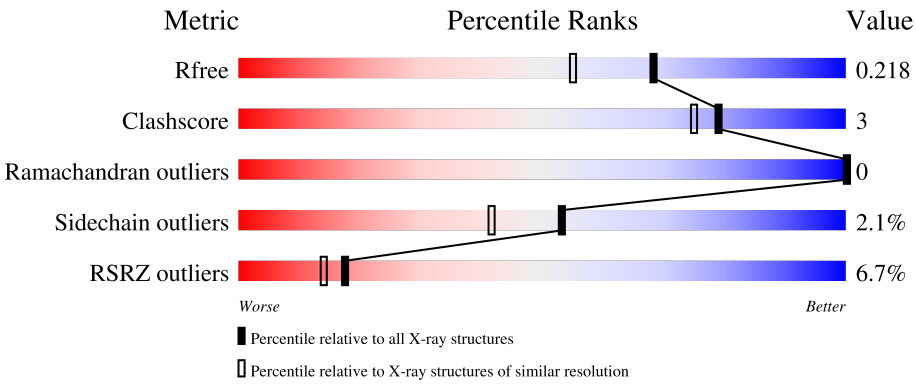
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

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}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (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 <=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	2%, 91%, 6%
1	C	331	3%, 88%, 10%
1	E	331	2%, 89%, 9%
2	B	176	20%, 87%, 9%
2	D	176	23%, 88%, 9%

Continued on next page...

Continued from previous page...

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

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MAN	H	5	-	-	-	X
6	NAG	A	401	-	-	-	X
6	NAG	C	401	-	-	-	X
6	NAG	E	401	-	-	-	X

2 Entry composition (i)

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
			Total	C	N	O	S			
1	A	323	2545	1604	440	490	11	0	5	0
1	C	323	2536	1598	438	489	11	0	4	0
1	E	325	2573	1622	442	498	11	0	8	0

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

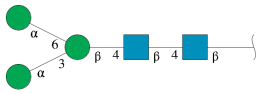
- Molecule 2 is a protein called Hemagglutinin HA2 chain.

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

- 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.

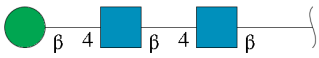
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	G	4	50	28	2	20	0	0	0

- 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.



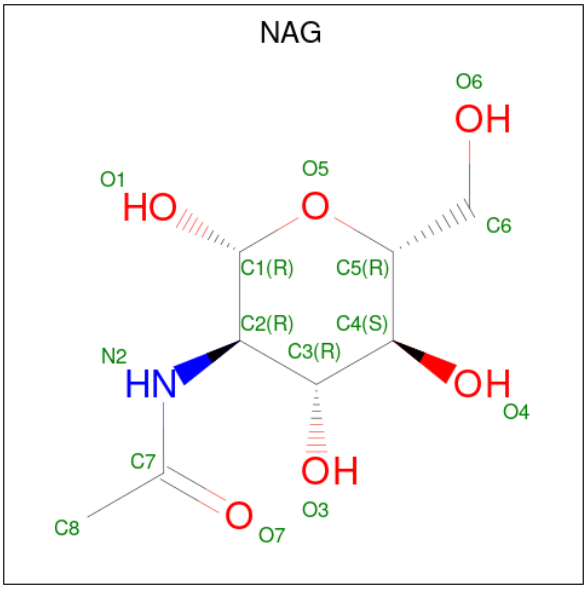
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	H	5	61	34	2	25	0	0	0

- 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
			Total	C	N	O			
5	I	3	39	22	2	15	0	0	0

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



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

- Molecule 7 is water.

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

Continued on next page...

Continued from previous page...

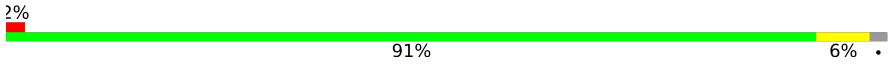
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	F	100	100	100	0	0

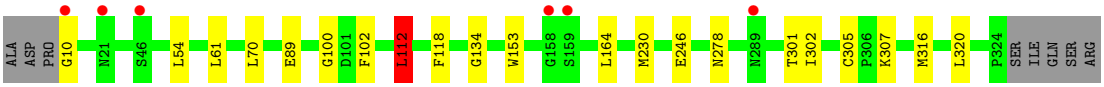
3 Residue-property plots (i)

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

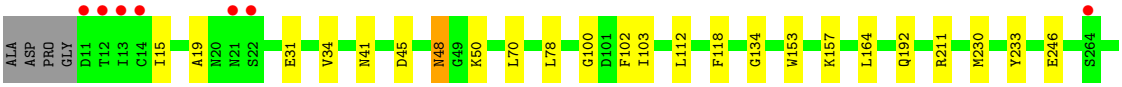
Chain A:





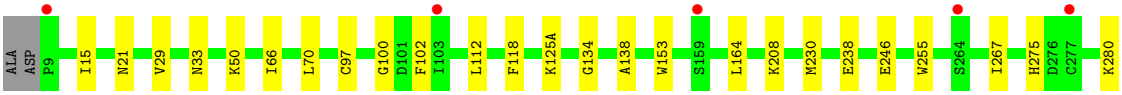
• Molecule 1: Hemagglutinin HA1 chain

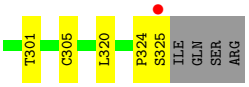
Chain C:



• Molecule 1: Hemagglutinin HA1 chain

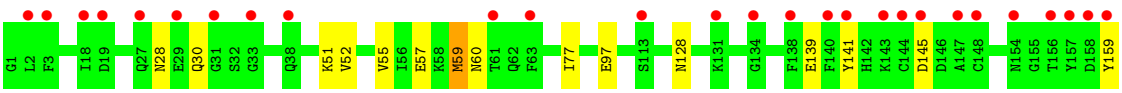
Chain E:

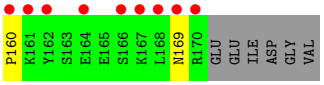




• Molecule 2: Hemagglutinin HA2 chain

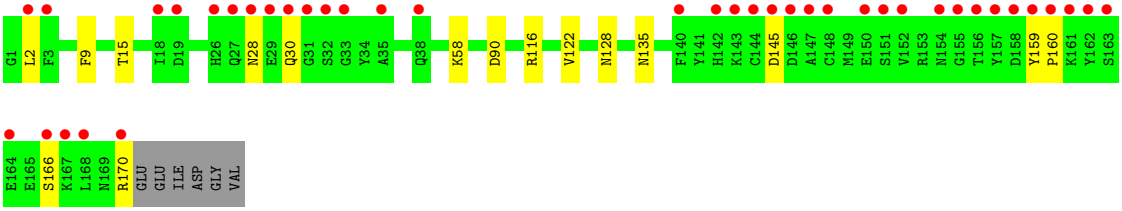
Chain B:





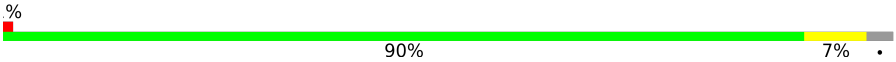
• Molecule 2: Hemagglutinin HA2 chain

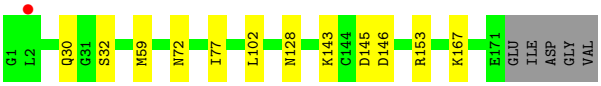
Chain D:



• Molecule 2: Hemagglutinin HA2 chain

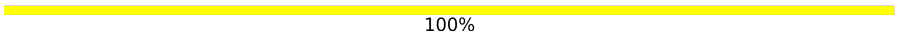
Chain F:





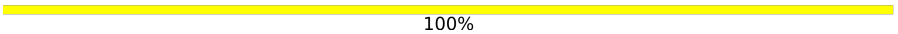
• Molecule 3: alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:



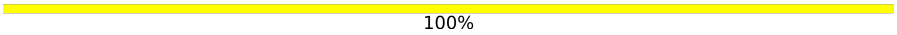
• Molecule 4: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain H:



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

Chain I:



4 Data and refinement statistics (i)

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 35.89 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.4 (35.90-1.80) 96.4 (35.89-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.47 (at 1.79Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.180 , 0.209 0.190 , 0.218	Depositor DCC
R_{free} test set	9530 reflections (5.04%)	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)	0.35 , 51.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	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.*

[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 (i)

5.1 Standard geometry (i)

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	0.54	0/2609	0.65	1/3551 (0.0%)
1	C	0.59	0/2600	0.67	0/3540
1	E	0.60	0/2638	0.67	0/3594
2	B	0.46	0/1414	0.55	0/1903
2	D	0.44	0/1412	0.53	0/1900
2	F	0.54	0/1426	0.64	1/1920 (0.1%)
All	All	0.55	0/12099	0.63	2/16408 (0.0%)

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	F	153	ARG	NE-CZ-NH2	-5.38	117.61	120.30
1	A	112	LEU	CA-CB-CG	-5.06	103.65	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2545	0	2468	12	0
1	C	2536	0	2458	19	0
1	E	2573	0	2494	20	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1387	0	1310	9	0
2	D	1385	0	1310	8	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	5	0
7	D	71	0	0	2	0
7	E	233	0	0	1	0
7	F	100	0	0	2	0
All	All	12942	0	11546	66	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 (66) 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: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: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
1:E:324:PRO:O	1:E:325:SER:C	2.45	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
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
1:C:157:LYS:NZ	7:C:647:HOH:O	2.47	0.48
2:D:116:ARG:HD3	7:D:261:HOH:O	2.15	0.47
1:C:134:GLY:HA3	1:C:153:TRP:HB3	1.96	0.46
2:F:167:LYS:NZ	7:F:390:HOH:O	2.48	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:F:143:LYS:NZ	7:F:325:HOH:O	2.48	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: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
1:C:45:ASP:C	1:C:297[A]:ILE:HD11	2.39	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:C:324:PRO:O	1:C:325:SER:C	2.56	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
1:E:97:CYS:HB2	1:E:138:ALA:O	2.19	0.42
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

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:31:GLU:HG3	1:C:34:VAL:CG2	2.50	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:316:MET:HE3	1:A:316:MET:HB2	1.99	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

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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 (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	282/284 (99%)	276 (98%)	6 (2%)	53	42
1	C	282/284 (99%)	278 (99%)	4 (1%)	67	59
1	E	287/284 (101%)	283 (99%)	4 (1%)	67	59
2	B	147/150 (98%)	142 (97%)	5 (3%)	37	22
2	D	147/150 (98%)	142 (97%)	5 (3%)	37	22
2	F	149/150 (99%)	145 (97%)	4 (3%)	44	31
All	All	1294/1302 (99%)	1266 (98%)	28 (2%)	53	39

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	102	PHE
1	A	112	LEU
1	A	118	PHE
1	A	278	ASN
1	A	307	LYS
1	A	320	LEU
2	B	28	ASN
2	B	57	GLU
2	B	59	MET
2	B	60	ASN
2	B	128	ASN
1	C	48	ASN
1	C	102	PHE
1	C	118	PHE
1	C	211	ARG
2	D	2	LEU
2	D	28	ASN
2	D	30	GLN
2	D	128	ASN
2	D	145	ASP
1	E	33	ASN
1	E	102	PHE
1	E	118	PHE
1	E	320	LEU
2	F	32	SER
2	F	72[A]	ASN
2	F	72[B]	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	128	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (26) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	41	ASN
1	A	191	GLN
1	A	197	ASN
1	A	250	ASN
1	A	275	HIS
2	B	28	ASN
2	B	125	GLN
2	B	128	ASN
2	B	129	ASN
1	C	41	ASN
1	C	48	ASN
1	C	191	GLN
1	C	250	ASN
2	D	28	ASN
2	D	30	GLN
2	D	128	ASN
2	D	129	ASN
1	E	191	GLN
1	E	197	ASN
1	E	250	ASN
1	E	275	HIS
2	F	30	GLN
2	F	43	ASN
2	F	125	GLN
2	F	128	ASN
2	F	129	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

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	1,3	14,14,15	0.75	1 (7%)	17,19,21	1.26	2 (11%)
3	NAG	G	2	3	14,14,15	0.71	1 (7%)	17,19,21	1.49	2 (11%)
3	BMA	G	3	3	11,11,12	1.27	2 (18%)	15,15,17	1.00	0
3	MAN	G	4	3	11,11,12	0.48	0	15,15,17	1.10	1 (6%)
4	NAG	H	1	1,4	14,14,15	0.62	0	17,19,21	1.29	1 (5%)
4	NAG	H	2	4	14,14,15	0.67	0	17,19,21	1.01	1 (5%)
4	BMA	H	3	4	11,11,12	1.65	2 (18%)	15,15,17	1.26	1 (6%)
4	MAN	H	4	4	11,11,12	0.57	0	15,15,17	1.18	1 (6%)
4	MAN	H	5	4	11,11,12	0.49	0	15,15,17	0.86	1 (6%)
5	NAG	I	1	1,5	14,14,15	0.81	1 (7%)	17,19,21	1.03	1 (5%)
5	NAG	I	2	5	14,14,15	0.73	0	17,19,21	1.07	1 (5%)
5	BMA	I	3	5	11,11,12	1.31	2 (18%)	15,15,17	1.05	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	1,3	-	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	1,4	-	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

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	I	1	1,5	-	0/6/23/26	0/1/1/1
5	NAG	I	2	5	-	0/6/23/26	0/1/1/1
5	BMA	I	3	5	-	0/2/19/22	0/1/1/1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	3	BMA	O5-C1	-3.37	1.38	1.43
3	G	3	BMA	C2-C3	2.94	1.56	1.52
4	H	3	BMA	C2-C3	2.54	1.56	1.52
5	I	3	BMA	O5-C5	2.40	1.48	1.43
3	G	1	NAG	C1-C2	2.29	1.55	1.52
5	I	1	NAG	O5-C1	-2.17	1.40	1.43
5	I	3	BMA	C4-C5	2.16	1.57	1.53
3	G	2	NAG	O5-C1	-2.16	1.40	1.43
3	G	3	BMA	C1-C2	2.07	1.56	1.52

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	H	1	NAG	O5-C1-C2	-3.65	105.52	111.29
3	G	2	NAG	C1-O5-C5	3.57	117.03	112.19
3	G	1	NAG	O5-C1-C2	-3.49	105.78	111.29
3	G	4	MAN	O5-C5-C6	2.95	111.82	107.20
4	H	5	MAN	O5-C5-C6	2.63	111.33	107.20
4	H	4	MAN	O5-C5-C6	2.49	111.11	107.20
4	H	2	NAG	O5-C5-C6	2.45	111.04	107.20
4	H	3	BMA	C1-O5-C5	2.36	115.39	112.19
5	I	3	BMA	C1-O5-C5	2.26	115.25	112.19
5	I	1	NAG	O5-C1-C2	-2.17	107.86	111.29
5	I	2	NAG	O5-C1-C2	-2.17	107.86	111.29
3	G	1	NAG	C1-C2-N2	2.16	114.18	110.49
3	G	2	NAG	O7-C7-C8	-2.15	118.06	122.06

There are no chirality outliers.

All (6) 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

Continued on next page...

Continued from previous page...

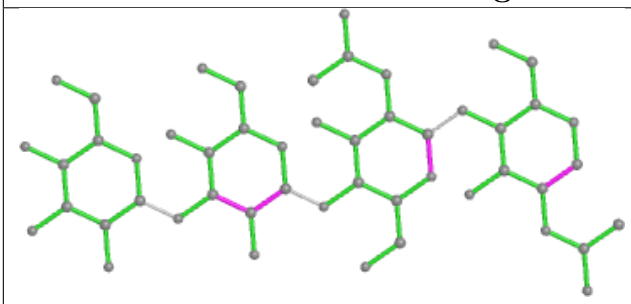
Mol	Chain	Res	Type	Atoms
3	G	2	NAG	C4-C5-C6-O6
4	H	5	MAN	C4-C5-C6-O6
4	H	5	MAN	O5-C5-C6-O6

There are no ring outliers.

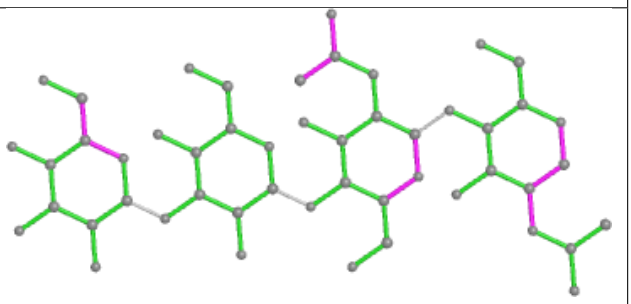
No monomer is involved in short contacts.

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

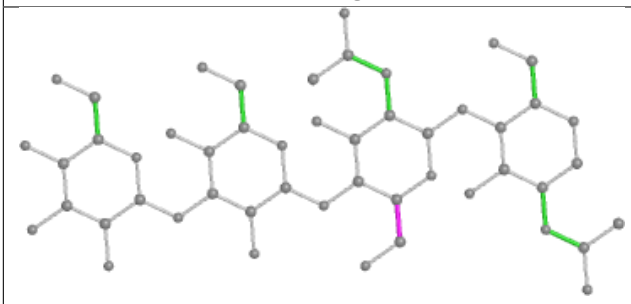
Oligosaccharide Chain G



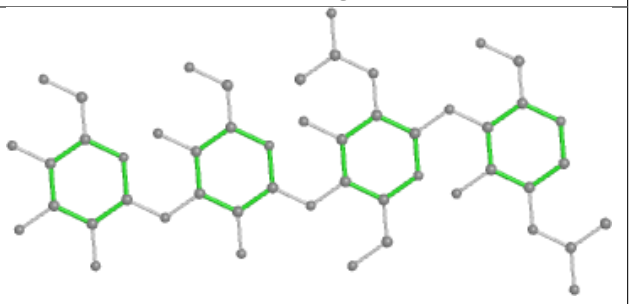
Bond lengths



Bond angles

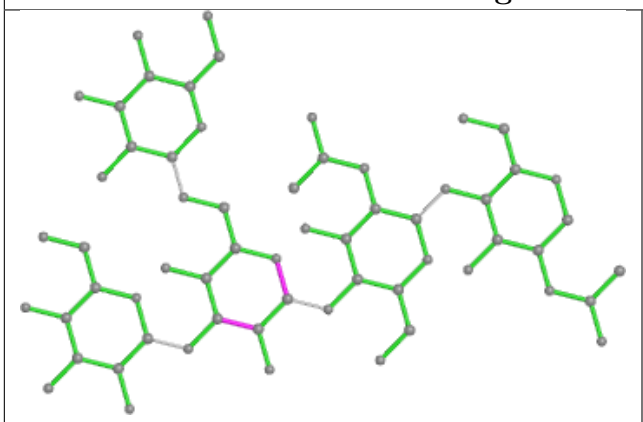


Torsions

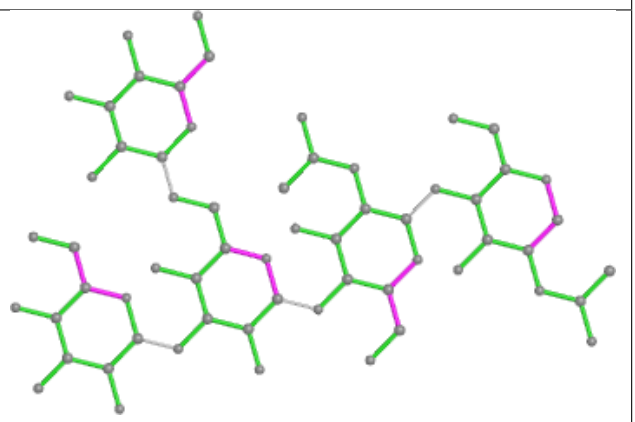


Rings

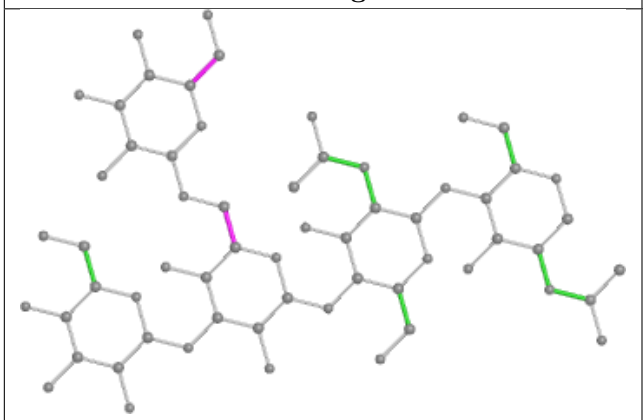
Oligosaccharide Chain H



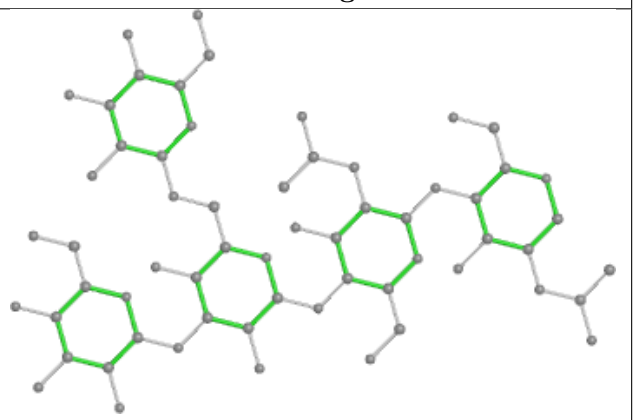
Bond lengths



Bond angles

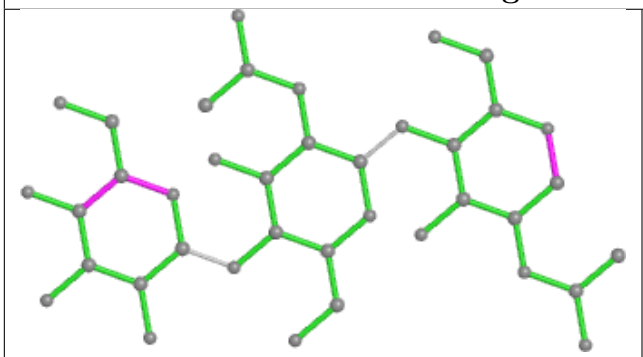


Torsions

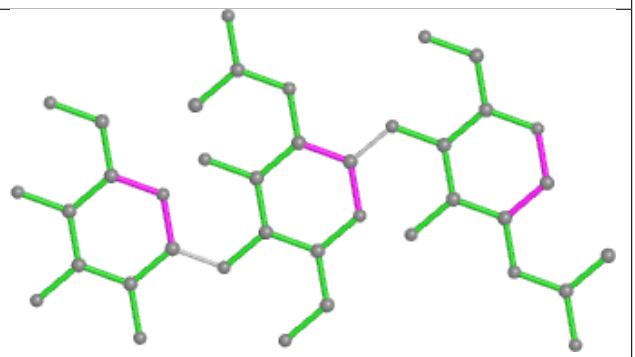


Rings

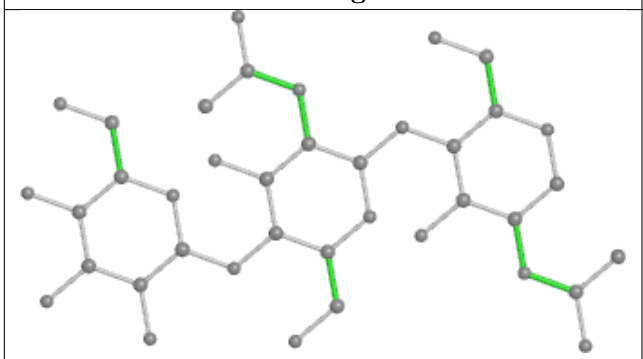
Oligosaccharide Chain I



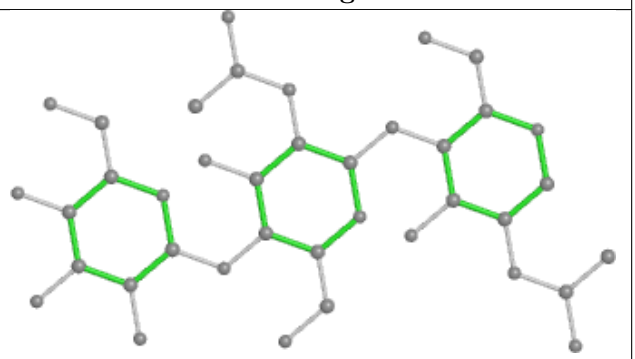
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

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	0.51	0	17,19,21	0.99	1 (5%)
6	NAG	A	401	1	14,14,15	0.54	0	17,19,21	1.06	1 (5%)
6	NAG	F	201	2	14,14,15	0.64	0	17,19,21	1.39	3 (17%)
6	NAG	C	407	1	14,14,15	0.51	0	17,19,21	0.81	0
6	NAG	C	401	1	14,14,15	0.62	0	17,19,21	1.30	3 (17%)

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

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	401	NAG	C1-O5-C5	3.58	117.04	112.19
6	F	201	NAG	O5-C5-C6	3.14	112.13	107.20
6	C	401	NAG	C1-O5-C5	3.09	116.38	112.19
6	F	201	NAG	C1-O5-C5	2.99	116.24	112.19
6	E	401	NAG	C1-O5-C5	2.76	115.94	112.19
6	C	401	NAG	O5-C1-C2	2.69	115.53	111.29
6	F	201	NAG	C3-C4-C5	-2.59	105.62	110.24

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	C	401	NAG	O5-C5-C6	2.12	110.53	107.20

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 (i)

6.1 Protein, DNA and RNA chains (i)

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%)	-0.09	6 (1%)	66	63	25, 44, 71, 97	0
1	C	323/331 (97%)	-0.05	10 (3%)	49	43	24, 43, 73, 103	0
1	E	325/331 (98%)	-0.05	6 (1%)	68	64	23, 41, 65, 90	0
2	B	170/176 (96%)	0.98	36 (21%)	0	0	26, 70, 105, 115	0
2	D	170/176 (96%)	1.29	40 (23%)	0	0	28, 69, 116, 129	0
2	F	171/176 (97%)	0.04	1 (0%)	89	87	27, 47, 70, 109	0
All	All	1482/1521 (97%)	0.22	99 (6%)	17	14	23, 46, 94, 129	0

All (99) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	168	LEU	11.7
2	B	168	LEU	10.5
2	D	144	CYS	9.3
2	D	148	CYS	8.2
2	D	140	PHE	7.2
2	D	143	LYS	6.6
2	D	156	THR	6.5
2	D	29	GLU	6.2
2	D	160	PRO	5.5
2	D	162	TYR	5.5
2	D	31	GLY	5.4
1	E	325	SER	5.1
2	D	33	GLY	5.0
2	B	157	TYR	4.9
1	C	325	SER	4.9
2	B	143	LYS	4.8
2	D	18	ILE	4.4
2	D	32	SER	4.3
2	D	147	ALA	4.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	167	LYS	4.3
2	B	167	LYS	4.2
2	D	28	ASN	4.2
2	D	146	ASP	4.2
2	B	144	CYS	3.9
2	B	29	GLU	3.9
2	B	169	ASN	3.9
2	B	156	THR	3.7
2	D	142	HIS	3.7
2	D	145	ASP	3.7
1	C	277	CYS	3.7
2	D	154	ASN	3.6
2	D	27	GLN	3.6
2	D	151	SER	3.6
2	D	163	SER	3.5
2	B	141	TYR	3.5
2	B	162	TYR	3.4
2	D	2	LEU	3.4
2	D	26	HIS	3.4
2	B	164	GLU	3.4
1	A	10	GLY	3.4
1	E	264	SER	3.3
2	B	161	LYS	3.3
2	D	30	GLN	3.3
2	B	38	GLN	3.2
1	C	12	THR	3.1
2	D	164	GLU	3.1
2	B	134	GLY	3.1
2	D	159	TYR	3.0
2	B	147	ALA	3.0
2	D	166	SER	2.9
2	B	140	PHE	2.8
2	B	148	CYS	2.8
2	B	145	ASP	2.8
2	B	159	TYR	2.8
1	E	277	CYS	2.8
2	B	160	PRO	2.8
2	D	157	TYR	2.7
2	D	161	LYS	2.7
2	B	27	GLN	2.7
2	B	33	GLY	2.7
2	B	170	ARG	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	19	ASP	2.6
1	A	21	ASN	2.6
2	B	158	ASP	2.6
2	B	18	ILE	2.5
1	C	11	ASP	2.5
2	B	131	LYS	2.5
1	E	9	PRO	2.5
2	D	3	PHE	2.5
2	B	61	THR	2.5
2	B	3	PHE	2.5
1	C	14	CYS	2.4
2	B	154	ASN	2.4
1	C	21	ASN	2.4
1	A	159	SER	2.4
1	C	324	PRO	2.4
2	B	166	SER	2.3
2	B	2	LEU	2.3
2	D	150	GLU	2.3
2	B	113	SER	2.2
2	D	152	VAL	2.2
2	F	2	LEU	2.2
2	D	158	ASP	2.2
2	B	138	PHE	2.2
2	B	63	PHE	2.2
2	D	170	ARG	2.2
1	C	13	ILE	2.2
2	D	38	GLN	2.1
1	A	158	GLY	2.1
1	A	289	ASN	2.1
2	B	31	GLY	2.1
2	D	155	GLY	2.1
1	C	22	SER	2.1
2	D	35	ALA	2.1
2	B	19	ASP	2.1
1	E	103[A]	ILE	2.0
1	A	46	SER	2.0
1	C	264	SER	2.0
1	E	159	SER	2.0

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

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

6.3 Carbohydrates (i)

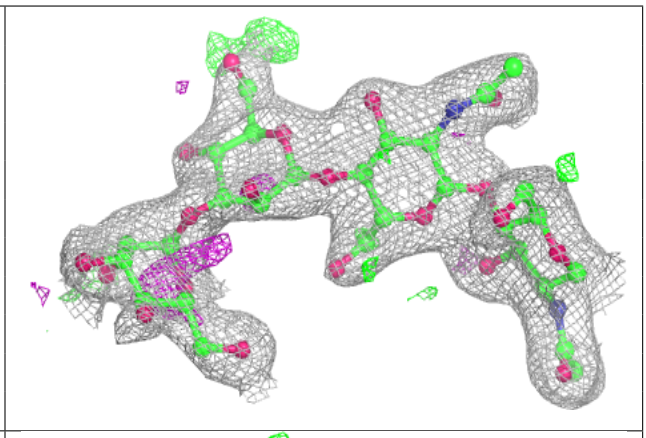
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

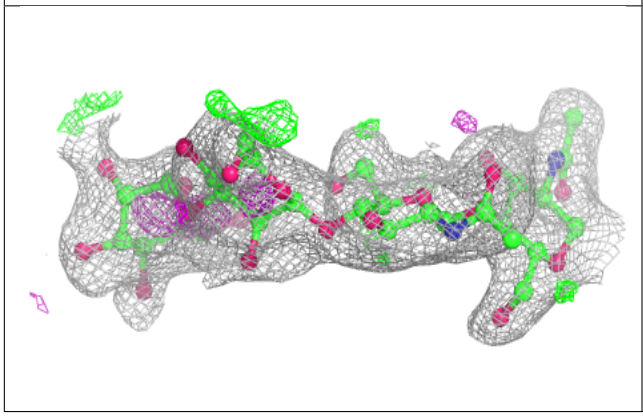
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	MAN	H	5	11/12	0.48	0.46	106,111,114,114	0
5	BMA	I	3	11/12	0.68	0.22	72,79,85,85	0
3	MAN	G	4	11/12	0.75	0.23	47,79,84,85	0
4	MAN	H	4	11/12	0.78	0.20	50,65,70,70	0
4	BMA	H	3	11/12	0.85	0.21	67,80,103,110	0
3	BMA	G	3	11/12	0.86	0.30	78,86,89,89	0
3	NAG	G	2	14/15	0.92	0.14	51,64,72,76	0
5	NAG	I	2	14/15	0.92	0.18	50,55,66,72	0
4	NAG	H	2	14/15	0.92	0.14	47,61,67,72	0
3	NAG	G	1	14/15	0.94	0.07	29,37,47,51	0
4	NAG	H	1	14/15	0.96	0.08	27,37,48,51	0
5	NAG	I	1	14/15	0.96	0.06	28,32,45,51	0

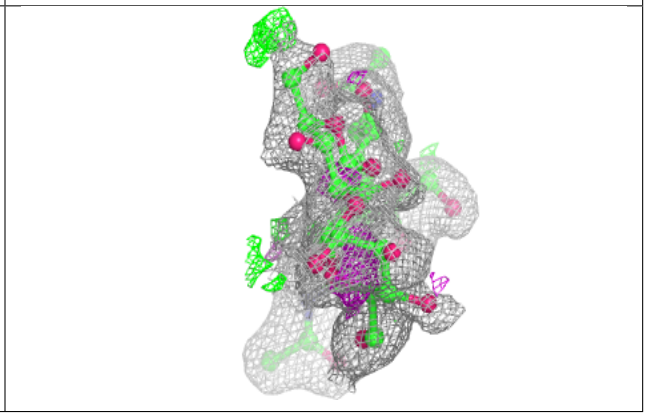
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

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)

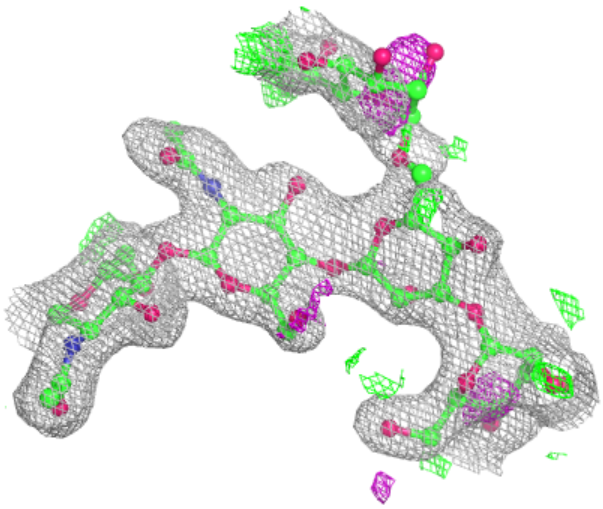


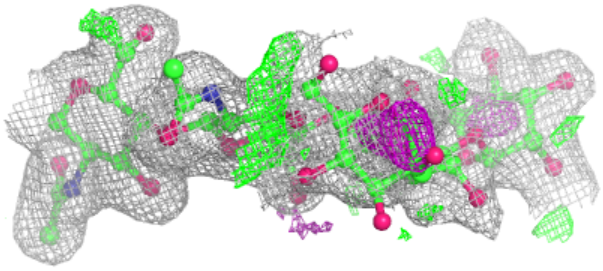


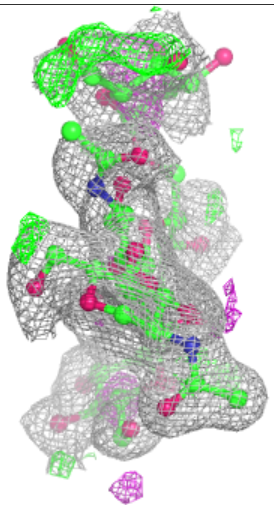


Electron density around Chain H:

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

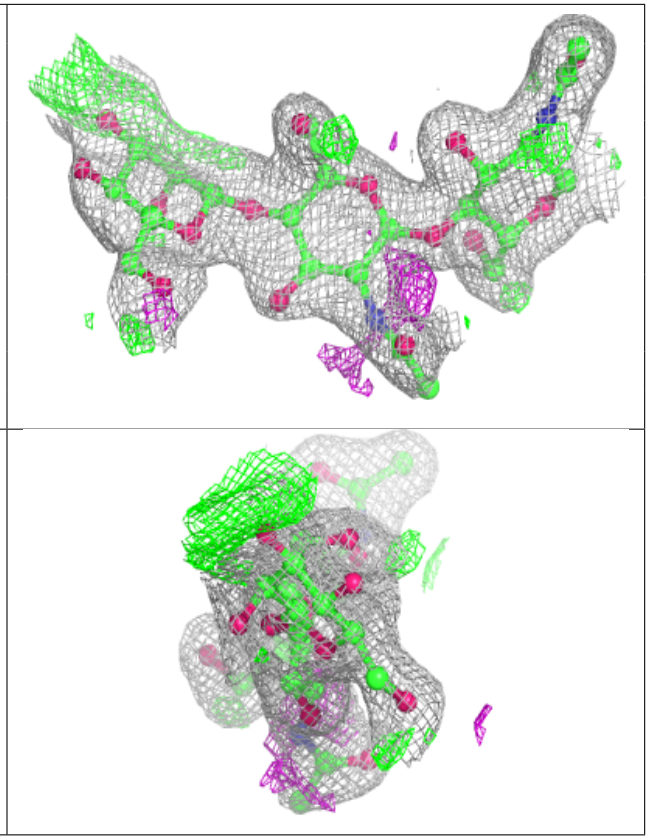






Electron density around Chain I:

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



6.4 Ligands (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
6	NAG	C	401	14/15	0.68	0.49	84,100,104,105	0
6	NAG	C	407	14/15	0.76	0.38	81,97,99,101	0
6	NAG	F	201	14/15	0.76	0.39	60,79,89,90	0
6	NAG	A	401	14/15	0.78	0.50	81,98,103,104	0
6	NAG	E	401	14/15	0.79	0.44	87,104,108,110	0

6.5 Other polymers (i)

There are no such residues in this entry.