Full wwPDB X-ray Structure Validation Report (i)

Jun 18, 2024 – 11:21 PM EDT

PDB ID	:	4BH0
Title	:	H5 (tyTy) Influenza Virus Haemagglutinin in Complex with Human Receptor Analogue 6'-SLN
Authors	:	Xiong, X.; Coombs, P.J.; Martin, S.R.; Liu, J.; Xiao, H.; McCauley, J.W.; Locher, K.; Walker, P.A.; Collins, P.J.; Kawaoka, Y.; Skehel, J.J.; Gamblin, S.J.
Deposited on	:	2013-03-29
Resolution	:	2.36 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:

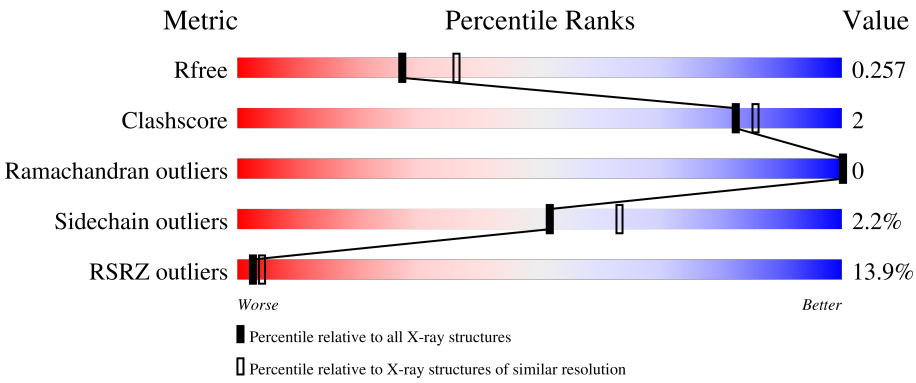
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.36 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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	327	4% 90% 7% • •
1	C	327	4% 91% 7% •
1	E	327	2% 90% 8% •
2	B	166	42% 83% • • 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	166	37% 84% 5% 11%
2	F	166	18% 82% 7% 11%
3	G	2	50% 50%
4	H	3	100%
4	I	3	33% 67%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 11922 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
			Total	C	N	O	S			
1	A	319	2526	1592	439	481	14	0	0	0
1	C	319	2519	1589	439	477	14	0	0	0
1	E	320	2523	1591	440	478	14	0	0	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	PRO	-	expression tag	UNP Q207Z6
A	323	ARG	-	expression tag	UNP Q207Z6
A	324	GLU	-	expression tag	UNP Q207Z6
A	325	THR	-	expression tag	UNP Q207Z6
A	326	ARG	-	expression tag	UNP Q207Z6
C	0	PRO	-	expression tag	UNP Q207Z6
C	323	ARG	-	expression tag	UNP Q207Z6
C	324	GLU	-	expression tag	UNP Q207Z6
C	325	THR	-	expression tag	UNP Q207Z6
C	326	ARG	-	expression tag	UNP Q207Z6
E	0	PRO	-	expression tag	UNP Q207Z6
E	323	ARG	-	expression tag	UNP Q207Z6
E	324	GLU	-	expression tag	UNP Q207Z6
E	325	THR	-	expression tag	UNP Q207Z6
E	326	ARG	-	expression tag	UNP Q207Z6

- Molecule 2 is a protein called HEMAGGLUTININ.

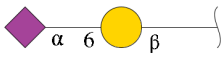
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	146	1131	698	199	226	8	0	0	0

Continued on next page...

Continued from previous page...

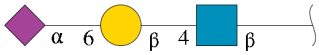
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	148	1159	718	203	230	8	0	0	0
2	F	148	1161	716	207	230	8	0	0	0

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galacto pyranose.



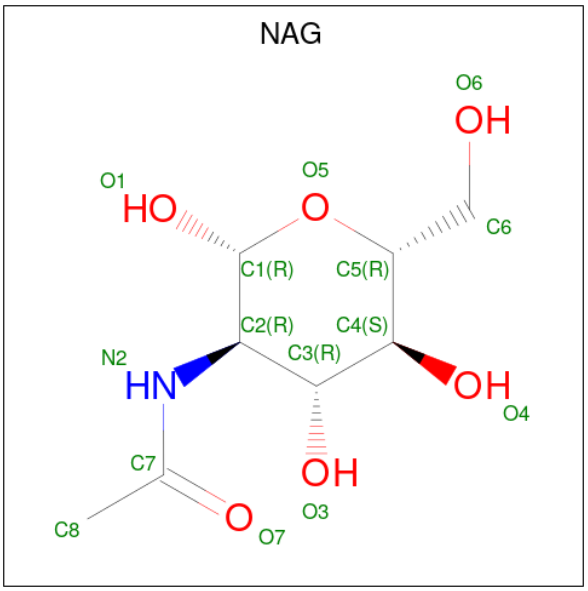
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	G	2	32	17	1	14	0	0	0

- Molecule 4 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galacto pyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



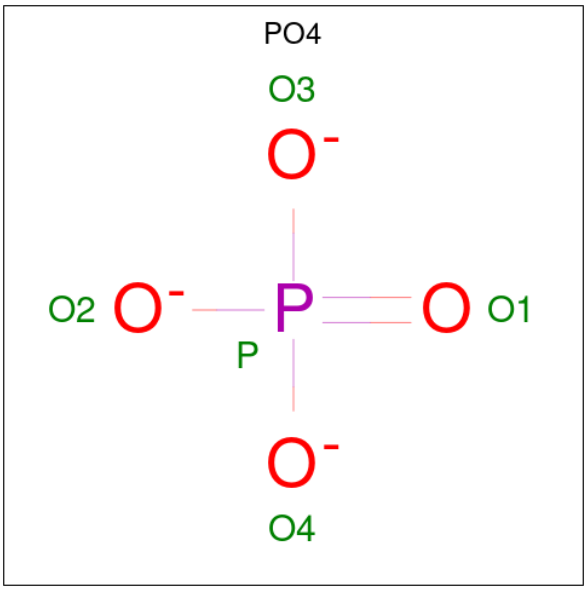
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	H	3	46	25	2	19	0	0	0
4	I	3	46	25	2	19	0	0	0

- Molecule 5 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		
5	A	1	14	8	1	5	0	0
5	C	1	14	8	1	5	0	0
5	E	1	14	8	1	5	0	0

- Molecule 6 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
6	A	1	5	4	1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
6	A	1	5	4	1	0	0
			Total	O	P		
6	A	1	5	4	1	0	0
			Total	O	P		
6	C	1	5	4	1	0	0
			Total	O	P		
6	E	1	5	4	1	0	0

- Molecule 7 is water.

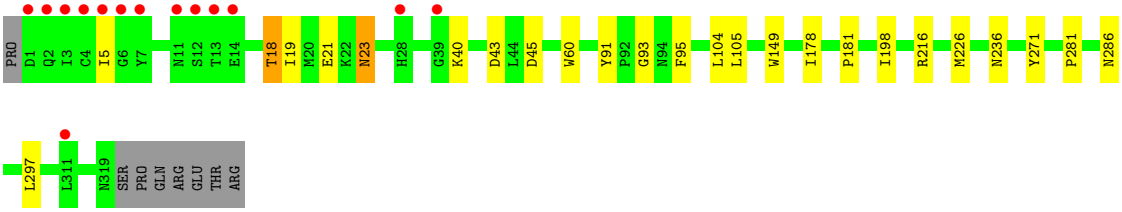
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	180	180	180	0	0
			Total	O		
7	B	53	53	53	0	0
			Total	O		
7	C	208	208	208	0	0
			Total	O		
7	D	44	44	44	0	0
			Total	O		
7	E	176	176	176	0	0
			Total	O		
7	F	51	51	51	0	0

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.

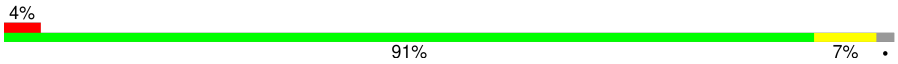
• Molecule 1: HEMAGGLUTININ

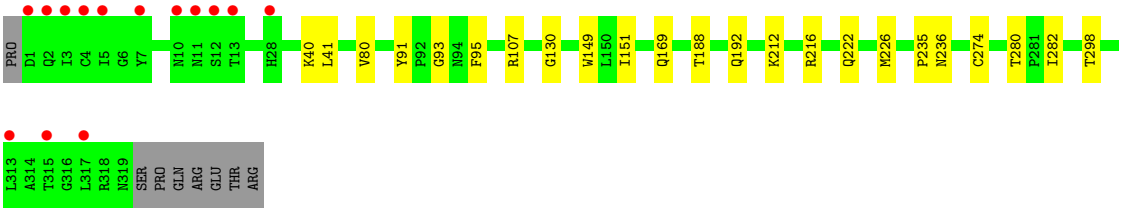
Chain A:



• Molecule 1: HEMAGGLUTININ

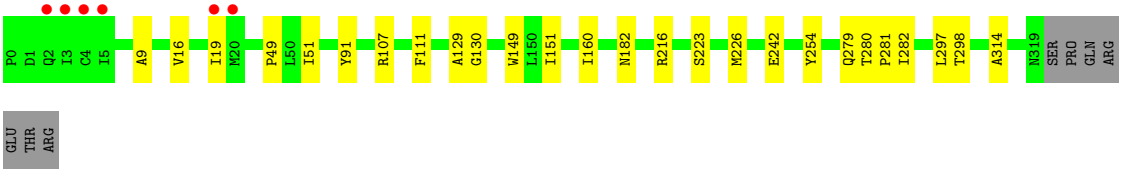
Chain C:





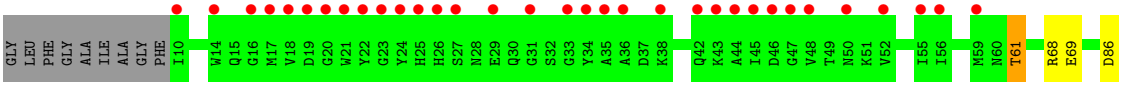
• Molecule 1: HEMAGGLUTININ

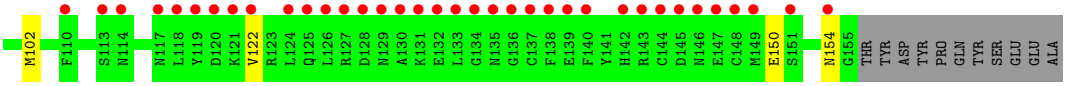
Chain E:



• Molecule 2: HEMAGGLUTININ

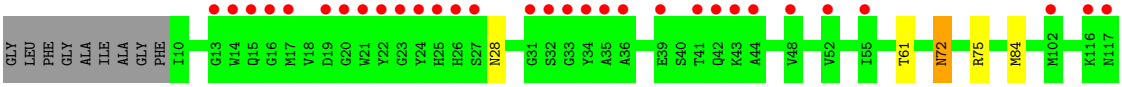
Chain B:

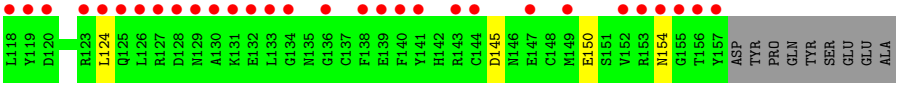




• Molecule 2: HEMAGGLUTININ

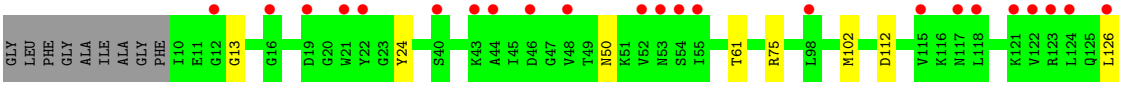
Chain D:

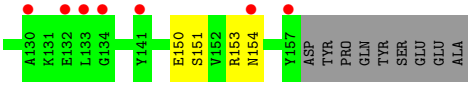




• Molecule 2: HEMAGGLUTININ

Chain F:





• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose

Chain G:

• Molecule 4: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain H:

• Molecule 4: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(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	70.35Å 228.30Å 71.27Å 90.00° 114.01° 90.00°	Depositor
Resolution (Å)	56.61 – 2.36 56.55 – 2.36	Depositor EDS
% Data completeness (in resolution range)	97.0 (56.61-2.36) 97.0 (56.55-2.36)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.88 (at 2.37Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.211 , 0.258 0.213 , 0.257	Depositor DCC
R_{free} test set	4094 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	40.9	Xtriage
Anisotropy	0.633	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 38.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	0.067 for l,-k,h	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	11922	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.53% 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: NAG, PO4, GAL, SIA

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	0.37	0/2586	0.57	0/3515
1	C	0.36	0/2579	0.58	0/3506
1	E	0.37	0/2584	0.58	2/3513 (0.1%)
2	B	0.31	0/1152	0.40	0/1556
2	D	0.30	0/1181	0.43	0/1594
2	F	0.32	0/1182	0.45	0/1594
All	All	0.35	0/11264	0.53	2/15278 (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})
1	E	216	ARG	NE-CZ-NH2	-6.15	117.22	120.30
1	E	216	ARG	NE-CZ-NH1	5.59	123.10	120.30

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	2526	0	2460	15	0
1	C	2519	0	2454	12	0
1	E	2523	0	2452	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1131	0	993	7	0
2	D	1159	0	1030	3	0
2	F	1161	0	1041	5	0
3	G	32	0	28	1	0
4	H	46	0	40	3	0
4	I	46	0	40	0	0
5	A	14	0	13	0	0
5	C	14	0	13	0	0
5	E	14	0	13	0	0
6	A	15	0	0	1	0
6	C	5	0	0	0	0
6	E	5	0	0	0	0
7	A	180	0	0	3	0
7	B	53	0	0	2	0
7	C	208	0	0	2	0
7	D	44	0	0	0	0
7	E	176	0	0	3	0
7	F	51	0	0	0	0
All	All	11922	0	10577	54	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 2.

All (54) 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:129:ALA:HB3	7:E:2089:HOH:O	1.76	0.84
1:C:280:THR:HG22	1:C:298:THR:HG22	1.68	0.75
1:A:286:ASN:ND2	7:A:2157:HOH:O	2.21	0.65
6:A:1325:PO4:O1	7:A:2180:HOH:O	2.14	0.64
1:C:169:GLN:NE2	7:C:2137:HOH:O	2.08	0.61
1:A:5:ILE:HD11	2:B:122:VAL:HG21	1.91	0.52
1:C:212:LYS:O	1:C:216:ARG:NH2	2.42	0.52
1:C:91:TYR:CD1	1:C:226:MET:HG2	2.45	0.52
1:A:93:GLY:HA3	1:A:226:MET:O	2.10	0.51
2:B:68:ARG:NH1	7:B:2021:HOH:O	2.26	0.51
1:E:19:ILE:HD11	2:F:102:MET:HG2	1.91	0.51
1:A:281:PRO:HD3	1:A:297:LEU:O	2.10	0.50
1:E:130:GLY:HA3	1:E:149:TRP:HB3	1.94	0.50
1:E:281:PRO:HD3	1:E:297:LEU:O	2.10	0.50
1:C:93:GLY:HA3	1:C:226:MET:O	2.14	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:72:ASN:HB2	2:D:75:ARG:NH2	2.29	0.48
1:C:188:THR:HG23	7:C:2146:HOH:O	2.13	0.48
1:E:151:ILE:HG12	7:E:2089:HOH:O	2.15	0.47
1:E:129:ALA:CB	7:E:2089:HOH:O	2.50	0.47
1:C:41:LEU:HD13	1:C:80:VAL:HG21	1.96	0.46
1:E:280:THR:HG22	1:E:298:THR:HG22	1.98	0.46
1:E:111:PHE:CD1	1:E:254:TYR:HB3	2.51	0.45
1:A:149:TRP:CZ2	3:G:2:SIA:H112	2.52	0.45
1:A:91:TYR:CD1	1:A:226:MET:HG2	2.52	0.45
2:B:61:THR:O	2:B:61:THR:OG1	2.32	0.45
1:E:16:VAL:HG21	1:E:314:ALA:HB2	1.99	0.44
2:F:24:TYR:CD1	2:F:153:ARG:HG2	2.52	0.44
4:H:1:NAG:H1	4:H:1:NAG:H82	1.99	0.44
1:C:235:PRO:O	1:C:236:ASN:HB2	2.17	0.44
1:E:279:GLN:NE2	1:E:280:THR:O	2.51	0.44
1:C:222:GLN:HE22	4:H:2:GAL:H5	1.83	0.44
1:E:49:PRO:HG2	1:E:51:ILE:HD11	2.00	0.44
1:C:151:ILE:HG21	4:H:3:SIA:H111	1.99	0.44
2:F:150:GLU:O	2:F:154:ASN:N	2.49	0.43
1:A:60:TRP:HZ3	1:A:105:LEU:HD21	1.82	0.43
1:E:9:ALA:HB2	2:F:13:GLY:HA3	2.00	0.43
1:A:178:ILE:HD12	1:A:198:ILE:HD12	2.01	0.43
1:A:18:THR:HG22	1:A:21:GLU:HB2	1.99	0.43
1:A:19:ILE:HD11	2:B:102:MET:CG	2.49	0.42
1:A:19:ILE:HD11	2:B:102:MET:HG2	2.00	0.42
2:D:28:ASN:HD21	2:D:145:ASP:HA	1.84	0.42
1:C:188:THR:HG22	1:C:192:GLN:O	2.20	0.42
1:E:91:TYR:CD1	1:E:226:MET:HG2	2.55	0.42
2:B:150:GLU:HG2	7:B:2052:HOH:O	2.20	0.41
1:E:160:ILE:O	1:E:242:GLU:HA	2.19	0.41
2:D:150:GLU:O	2:D:154:ASN:N	2.53	0.41
1:E:182:ASN:OD1	1:E:223:SER:HB3	2.19	0.41
1:A:43:ASP:OD1	1:A:271:TYR:OH	2.30	0.41
1:A:181:PRO:O	1:A:216:ARG:NH2	2.53	0.41
1:C:130:GLY:HA3	1:C:149:TRP:HB3	2.03	0.41
1:A:23:ASN:N	1:A:23:ASN:HD22	2.19	0.41
1:A:104:LEU:HD11	7:A:2143:HOH:O	2.20	0.41
1:E:19:ILE:HD11	2:F:102:MET:CG	2.50	0.41
2:B:150:GLU:O	2:B:154:ASN:N	2.50	0.40

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	317/327 (97%)	303 (96%)	14 (4%)	0	100	100
1	C	317/327 (97%)	301 (95%)	16 (5%)	0	100	100
1	E	318/327 (97%)	307 (96%)	11 (4%)	0	100	100
2	B	144/166 (87%)	139 (96%)	5 (4%)	0	100	100
2	D	146/166 (88%)	139 (95%)	7 (5%)	0	100	100
2	F	146/166 (88%)	145 (99%)	1 (1%)	0	100	100
All	All	1388/1479 (94%)	1334 (96%)	54 (4%)	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	284/293 (97%)	278 (98%)	6 (2%)	53	65
1	C	282/293 (96%)	277 (98%)	5 (2%)	59	70
1	E	282/293 (96%)	280 (99%)	2 (1%)	84	91
2	B	112/141 (79%)	109 (97%)	3 (3%)	44	55
2	D	116/141 (82%)	112 (97%)	4 (3%)	37	46
2	F	117/141 (83%)	111 (95%)	6 (5%)	24	27
All	All	1193/1302 (92%)	1167 (98%)	26 (2%)	52	63

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	THR
1	A	23	ASN
1	A	40	LYS
1	A	45	ASP
1	A	95	PHE
1	A	236	ASN
2	B	61	THR
2	B	69	GLU
2	B	86	ASP
1	C	40	LYS
1	C	95	PHE
1	C	107	ARG
1	C	274	CYS
1	C	282	ILE
2	D	61	THR
2	D	72	ASN
2	D	84	MET
2	D	124	LEU
1	E	107	ARG
1	E	282	ILE
2	F	50	ASN
2	F	61	THR
2	F	75	ARG
2	F	112	ASP
2	F	126	LEU
2	F	151	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	23	ASN
1	A	279	GLN
2	B	50	ASN
2	B	125	GLN
1	C	2	GLN
1	C	240	ASN
2	D	28	ASN
2	D	50	ASN
2	D	72	ASN
1	E	2	GLN
1	E	23	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	169	GLN
2	F	50	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)

8 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	GAL	G	1	3	12,12,12	0.57	0	17,17,17	1.02	1 (5%)
3	SIA	G	2	3	20,20,21	0.58	0	21,28,31	1.42	4 (19%)
4	NAG	H	1	4	15,15,15	0.67	0	21,21,21	1.54	2 (9%)
4	GAL	H	2	4	11,11,12	0.66	0	15,15,17	1.19	2 (13%)
4	SIA	H	3	4	20,20,21	0.58	0	21,28,31	1.27	2 (9%)
4	NAG	I	1	4	15,15,15	0.53	0	21,21,21	1.11	1 (4%)
4	GAL	I	2	4	11,11,12	0.49	0	15,15,17	1.08	0
4	SIA	I	3	4	20,20,21	0.64	0	21,28,31	1.29	4 (19%)

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	GAL	G	1	3	-	0/2/22/22	0/1/1/1
3	SIA	G	2	3	-	4/18/34/38	0/1/1/1
4	NAG	H	1	4	-	6/6/26/26	0/1/1/1
4	GAL	H	2	4	-	1/2/19/22	0/1/1/1
4	SIA	H	3	4	-	0/18/34/38	0/1/1/1
4	NAG	I	1	4	-	2/6/26/26	0/1/1/1
4	GAL	I	2	4	-	0/2/19/22	0/1/1/1
4	SIA	I	3	4	-	0/18/34/38	0/1/1/1

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	H	1	NAG	C2-N2-C7	4.42	133.46	123.11
4	H	3	SIA	C6-C5-N5	-3.73	104.95	110.91
3	G	2	SIA	C6-C5-N5	-3.69	105.02	110.91
4	H	1	NAG	C8-C7-N2	3.35	121.67	116.12
3	G	2	SIA	O1B-C1-C2	3.08	120.71	112.71
4	H	2	GAL	C1-C2-C3	3.00	114.02	109.64
4	I	3	SIA	C6-C5-N5	-2.71	106.58	110.91
3	G	2	SIA	O6-C2-C1	2.62	112.65	107.72
4	I	3	SIA	O1B-C1-C2	2.54	119.32	112.71
4	H	3	SIA	O1B-C1-C2	2.40	118.96	112.71
4	H	2	GAL	O5-C5-C4	-2.40	104.99	110.83
3	G	2	SIA	O1A-C1-C2	-2.38	117.71	122.85
4	I	3	SIA	O6-C2-C1	2.36	112.17	107.72
3	G	1	GAL	O5-C5-C6	2.33	112.22	106.44
4	I	1	NAG	C1-O5-C5	2.32	118.14	113.65
4	I	3	SIA	C4-C5-N5	-2.00	106.49	110.44

There are no chirality outliers.

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	2	SIA	C6-C7-C8-C9
3	G	2	SIA	O7-C7-C8-C9
4	H	1	NAG	O5-C5-C6-O6
3	G	2	SIA	C6-C7-C8-O8
4	H	1	NAG	C4-C5-C6-O6
3	G	2	SIA	O7-C7-C8-O8
4	H	1	NAG	C8-C7-N2-C2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	H	1	NAG	O7-C7-N2-C2
4	H	1	NAG	C1-C2-N2-C7
4	I	1	NAG	C3-C2-N2-C7
4	I	1	NAG	C1-C2-N2-C7
4	H	2	GAL	O5-C5-C6-O6
4	H	1	NAG	C3-C2-N2-C7

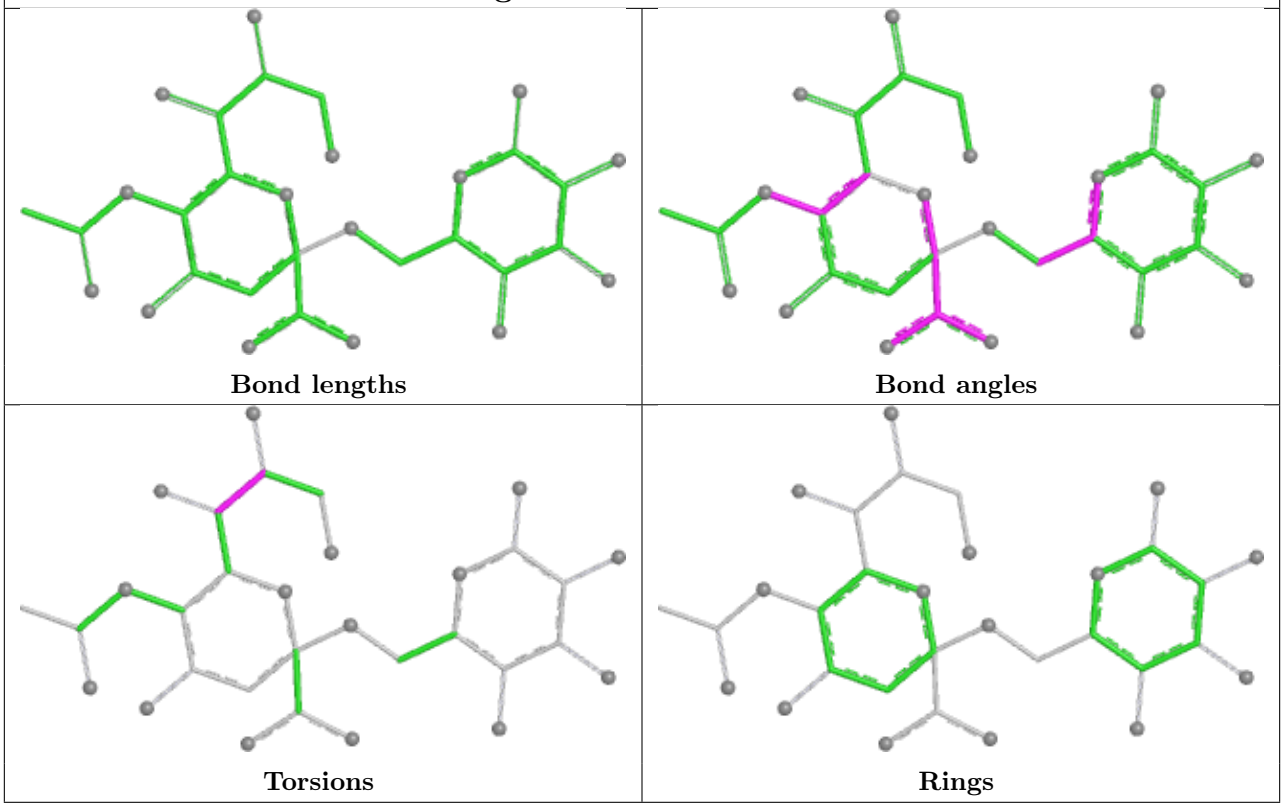
There are no ring outliers.

4 monomers are involved in 4 short contacts:

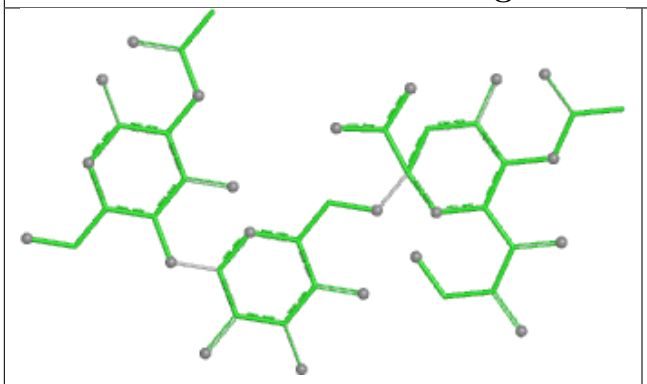
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	2	GAL	1	0
4	H	1	NAG	1	0
3	G	2	SIA	1	0
4	H	3	SIA	1	0

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

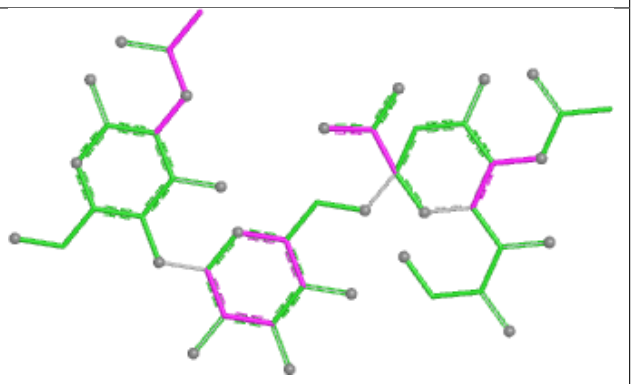
Oligosaccharide Chain G



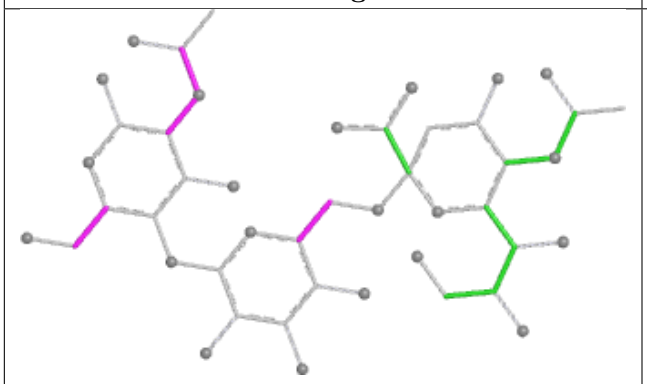
Oligosaccharide Chain H



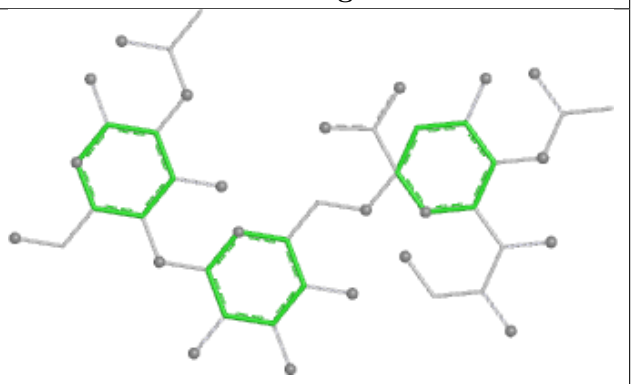
Bond lengths



Bond angles

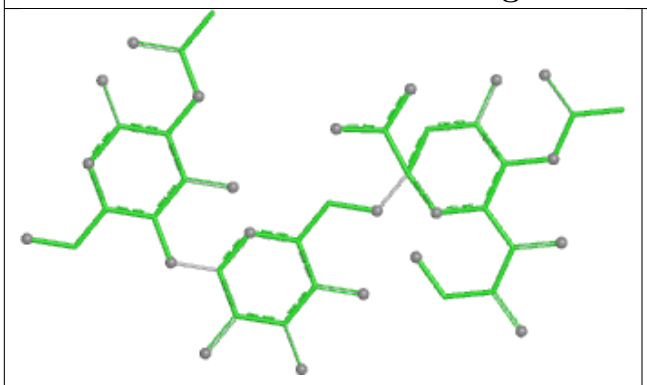


Torsions

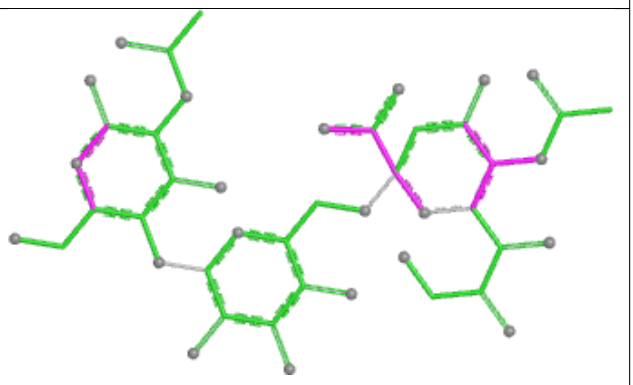


Rings

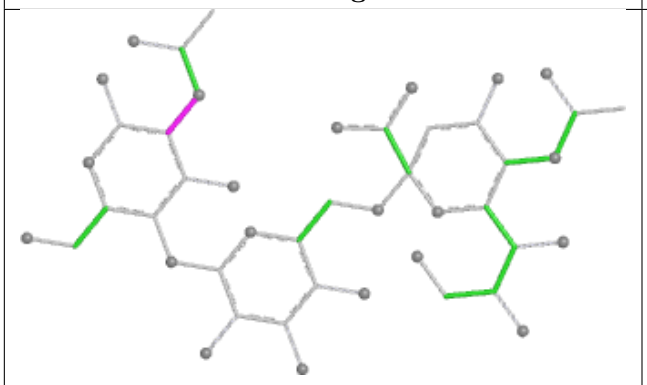
Oligosaccharide Chain I



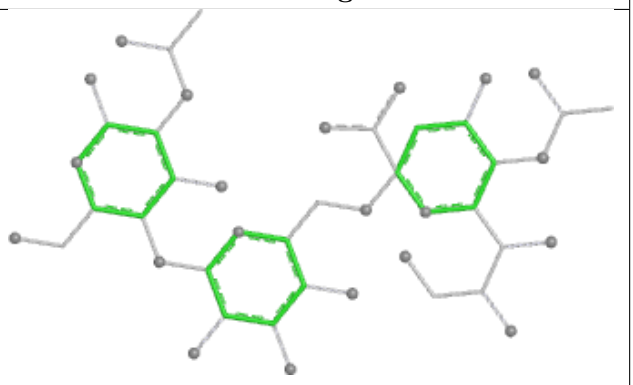
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

8 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	PO4	C	1324	-	4,4,4	0.91	0	6,6,6	0.39	0
6	PO4	A	1325	-	4,4,4	0.90	0	6,6,6	0.78	0
5	NAG	C	1320	1	14,14,15	0.45	0	17,19,21	1.15	1 (5%)
6	PO4	A	1323	-	4,4,4	0.91	0	6,6,6	0.59	0
6	PO4	A	1324	-	4,4,4	0.91	0	6,6,6	0.61	0
6	PO4	E	1324	-	4,4,4	0.85	0	6,6,6	0.53	0
5	NAG	E	1320	1	14,14,15	0.39	0	17,19,21	1.06	1 (5%)
5	NAG	A	1320	1	14,14,15	0.51	0	17,19,21	1.42	2 (11%)

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	E	1320	1	-	0/6/23/26	0/1/1/1
5	NAG	A	1320	1	-	0/6/23/26	0/1/1/1
5	NAG	C	1320	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	1320	NAG	C1-O5-C5	4.31	117.97	112.19
5	E	1320	NAG	C1-O5-C5	3.58	116.99	112.19
5	C	1320	NAG	C1-O5-C5	3.37	116.70	112.19
5	A	1320	NAG	C1-C2-N2	2.21	113.91	110.43

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1325	PO4	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	319/327 (97%)	0.21	14 (4%)	34	46	22, 42, 105, 181	0
1	C	319/327 (97%)	0.15	14 (4%)	34	46	23, 41, 95, 174	0
1	E	320/327 (97%)	0.10	6 (1%)	66	76	19, 43, 86, 102	0
2	B	146/166 (87%)	2.36	69 (47%)	0	0	22, 118, 168, 186	0
2	D	148/166 (89%)	2.17	61 (41%)	0	0	24, 116, 172, 201	0
2	F	148/166 (89%)	1.08	30 (20%)	1	1	24, 89, 120, 127	0
All	All	1400/1479 (94%)	0.70	194 (13%)	2	4	19, 49, 164, 201	0

All (194) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	20	GLY	17.5
1	C	2	GLN	11.3
2	B	138	PHE	11.1
2	D	152	VAL	10.5
1	C	1	ASP	9.3
2	D	140	PHE	9.2
2	B	22	TYR	9.0
2	B	36	ALA	8.9
2	D	33	GLY	8.7
1	A	5	ILE	8.6
2	D	21	TRP	8.6
2	B	128	ASP	8.5
2	D	138	PHE	8.4
2	F	21	TRP	8.3
2	B	130	ALA	8.2
1	A	3	ILE	8.1
2	B	26	HIS	7.9
2	B	134	GLY	7.9
2	D	156	THR	7.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	22	TYR	7.6
2	B	133	LEU	7.4
2	B	124	LEU	7.4
2	D	153	ARG	7.1
2	B	24	TYR	7.0
1	A	2	GLN	6.9
2	B	144	CYS	6.9
2	B	132	GLU	6.8
2	B	21	TRP	6.7
1	C	5	ILE	6.6
2	B	23	GLY	6.3
2	B	16	GLY	6.1
2	D	14	TRP	6.1
2	B	127	ARG	6.1
2	B	137	CYS	6.1
2	D	131	LYS	6.0
1	C	3	ILE	5.9
2	B	17	MET	5.9
2	D	24	TYR	5.8
2	D	125	GLN	5.8
2	B	18	VAL	5.7
2	B	19	ASP	5.5
2	B	121	LYS	5.4
2	D	27	SER	5.3
2	F	124	LEU	5.3
2	D	139	GLU	5.3
2	B	119	TYR	5.2
2	B	129	ASN	5.2
2	F	118	LEU	5.2
2	F	130	ALA	5.2
2	B	140	PHE	5.1
2	D	144	CYS	5.0
2	F	126	LEU	5.0
2	D	134	GLY	4.9
2	D	17	MET	4.9
2	F	141	TYR	4.8
2	D	154	ASN	4.8
1	C	4	CYS	4.7
2	D	52	VAL	4.6
2	B	31	GLY	4.6
2	B	46	ASP	4.5
2	D	155	GLY	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	33	GLY	4.5
2	D	141	TYR	4.5
2	B	149	MET	4.4
2	B	43	LYS	4.4
2	B	148	CYS	4.3
2	F	44	ALA	4.2
2	F	133	LEU	4.2
2	F	48	VAL	4.2
2	B	27	SER	4.2
2	D	15	GLN	4.1
2	B	25	HIS	4.1
1	A	13	THR	4.1
1	A	11	ASN	4.0
2	F	132	GLU	4.0
1	A	4	CYS	4.0
2	B	10	ILE	4.0
2	D	124	LEU	4.0
2	B	139	GLU	4.0
2	B	118	LEU	3.9
2	B	110	PHE	3.9
2	B	154	ASN	3.9
2	D	136	GLY	3.8
2	D	127	ARG	3.8
2	D	130	ALA	3.7
2	F	40	SER	3.7
1	C	317	LEU	3.6
2	D	157	TYR	3.6
2	F	122	VAL	3.6
2	D	32	SER	3.6
2	D	118	LEU	3.5
2	F	22	TYR	3.5
2	B	35	ALA	3.5
2	D	16	GLY	3.5
2	D	19	ASP	3.5
2	B	136	GLY	3.5
2	D	126	LEU	3.5
2	B	55	ILE	3.5
2	B	117	ASN	3.4
2	B	20	GLY	3.4
2	D	133	LEU	3.4
2	D	23	GLY	3.4
1	C	11	ASN	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1	ASP	3.3
2	F	54	SER	3.3
2	B	131	LYS	3.3
2	B	142	HIS	3.3
2	B	143	ARG	3.3
2	F	121	LYS	3.3
2	B	44	ALA	3.2
2	B	114	ASN	3.2
2	D	44	ALA	3.2
2	B	122	VAL	3.2
2	B	113	SER	3.2
2	B	147	GLU	3.2
2	D	26	HIS	3.2
1	A	12	SER	3.1
2	F	55	ILE	3.1
2	B	38	LYS	3.1
2	D	132	GLU	3.0
2	B	34	TYR	3.0
2	D	128	ASP	3.0
2	D	149	MET	3.0
2	B	126	LEU	2.9
1	E	3	ILE	2.9
2	D	55	ILE	2.9
2	D	143	ARG	2.9
1	A	7	TYR	2.9
2	B	151	SER	2.9
2	B	45	ILE	2.8
2	D	31	GLY	2.8
2	B	59	MET	2.8
2	D	48	VAL	2.8
2	D	25	HIS	2.8
2	F	134	GLY	2.8
1	E	5	ILE	2.8
2	F	98	LEU	2.7
2	D	39	GLU	2.7
2	D	116	LYS	2.7
2	F	154	ASN	2.7
2	F	157	TYR	2.7
2	D	13	GLY	2.7
2	D	123	ARG	2.7
1	C	315	THR	2.7
2	F	12	GLY	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	16	GLY	2.6
2	B	29	GLU	2.6
1	C	10	ASN	2.6
2	F	53	ASN	2.6
2	D	34	TYR	2.6
2	B	47	GLY	2.6
1	C	7	TYR	2.6
2	B	125	GLN	2.6
2	B	48	VAL	2.5
2	F	19	ASP	2.5
1	A	14	GLU	2.5
1	E	19	ILE	2.5
1	A	39	GLY	2.5
2	F	43	LYS	2.5
1	E	2	GLN	2.5
2	B	50	ASN	2.4
1	C	12	SER	2.4
2	D	43	LYS	2.4
2	D	36	ALA	2.4
2	D	41	THR	2.4
2	D	129	ASN	2.4
2	B	145	ASP	2.3
2	F	52	VAL	2.3
2	F	123	ARG	2.3
2	D	117	ASN	2.3
2	D	120	ASP	2.3
2	D	35	ALA	2.3
1	E	4	CYS	2.2
2	F	117	ASN	2.2
2	D	119	TYR	2.2
2	B	135	ASN	2.2
2	D	42	GLN	2.2
1	E	20	MET	2.2
2	B	52	VAL	2.2
2	D	147	GLU	2.2
2	F	46	ASP	2.2
2	B	146	ASN	2.2
2	B	56	ILE	2.2
2	B	14	TRP	2.2
1	A	28	HIS	2.1
1	C	313	LEU	2.1
2	B	120	ASP	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	42	GLN	2.1
1	C	28	HIS	2.1
2	D	102	MET	2.1
1	C	13	THR	2.1
1	A	311	LEU	2.1
1	A	6	GLY	2.1
2	F	115	VAL	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)

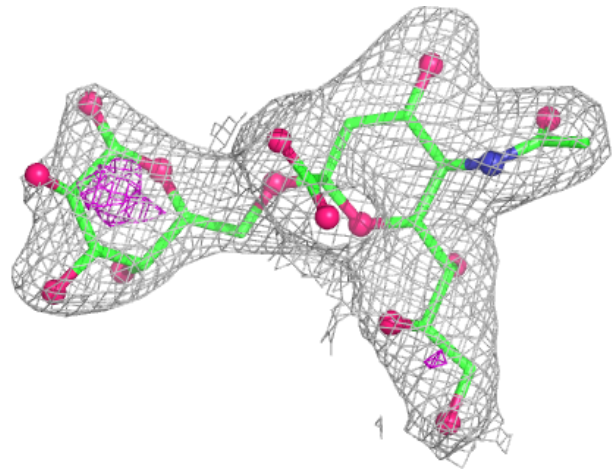
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

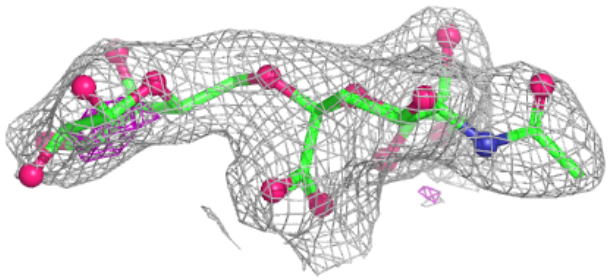
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	I	1	15/15	0.85	0.26	55,66,69,70	0
4	GAL	H	2	11/12	0.88	0.12	44,56,61,62	0
4	NAG	H	1	15/15	0.88	0.20	68,75,76,76	0
3	GAL	G	1	12/12	0.91	0.31	66,75,77,78	0
3	SIA	G	2	20/21	0.92	0.14	51,58,63,63	0
4	GAL	I	2	11/12	0.93	0.09	33,48,53,55	0
4	SIA	H	3	20/21	0.97	0.15	34,36,42,44	0
4	SIA	I	3	20/21	0.98	0.12	24,28,31,32	0

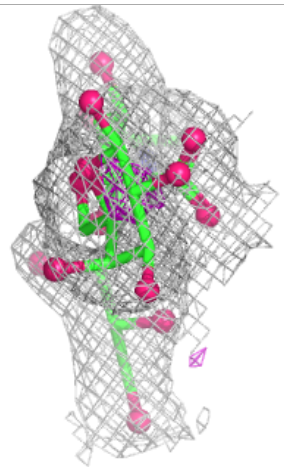
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

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)

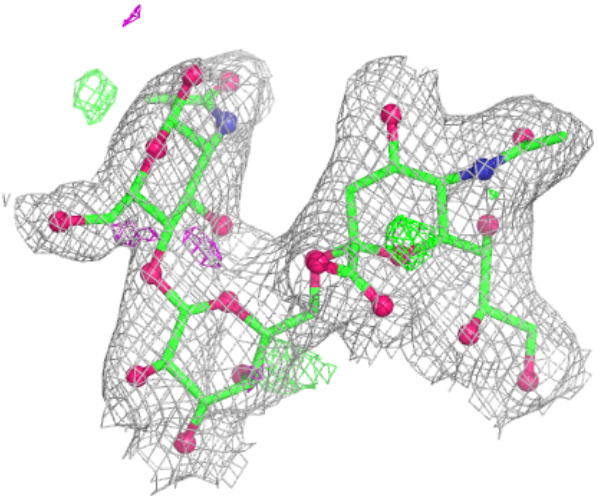


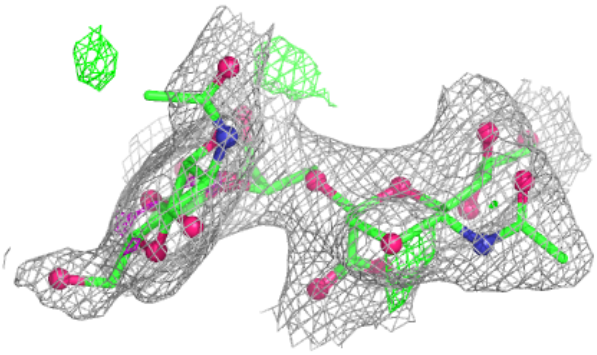


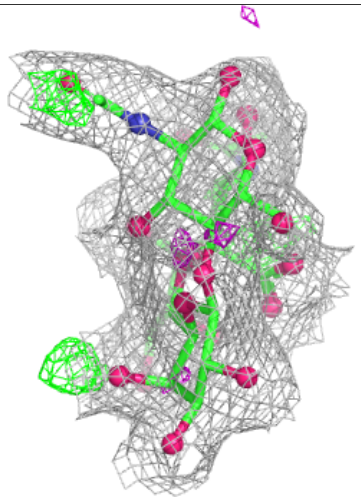


Electron density around Chain H:

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

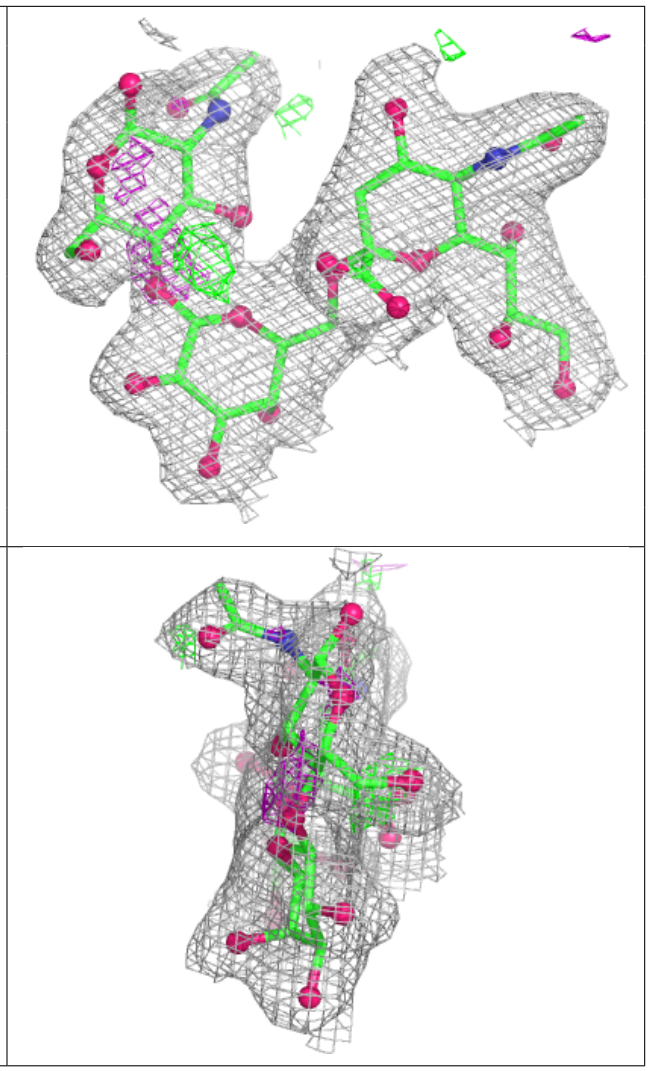






Electron density around Chain I:

$2mF_o$-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
5	NAG	A	1320	14/15	0.94	0.10	31,35,41,43	0
5	NAG	C	1320	14/15	0.94	0.11	37,41,44,48	0
5	NAG	E	1320	14/15	0.94	0.09	39,42,51,52	0
6	PO4	E	1324	5/5	0.94	0.33	67,70,72,73	0
6	PO4	A	1324	5/5	0.95	0.14	60,61,66,68	0
6	PO4	A	1325	5/5	0.97	0.17	65,67,70,70	0
6	PO4	C	1324	5/5	0.97	0.17	65,66,69,72	0
6	PO4	A	1323	5/5	0.97	0.20	71,73,75,76	0

6.5 Other polymers (i)

There are no such residues in this entry.