Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 09:44 AM BST

PDB ID	:	6CNV
Title	:	INFLUENZA B/BRISBANE HEMAGGLUTININ FAB CR9115 SD84H COMPLEX
Authors	:	Luo, J.; Obmolova, G.
Deposited on	:	2018-03-09
Resolution	:	4.10 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 following versions of software and data (see references ⓘ) were used in the production of this report:

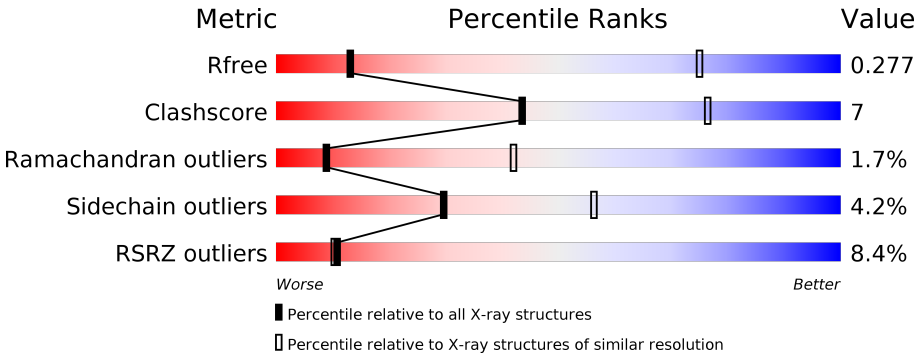
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 4.10 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	1193 (4.50-3.70)
Clashscore	141614	1003 (4.44-3.76)
Ramachandran outliers	138981	1005 (4.48-3.72)
Sidechain outliers	138945	1199 (4.50-3.70)
RSRZ outliers	127900	1034 (4.50-3.70)

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 on 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	347	% 81% 17% •
2	B	220	% 52% 14% • 32%
3	C	116	87% 12% •
4	L	216	23% 83% 13% •
5	H	230	14% 63% 25% 5% 7%
6	D	2	50% 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	E	2	50% 50%
6	F	2	100%
6	G	2	50% 50%

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
6	NAG	F	2	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7901 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	346	Total	C	N	O	S	0	0	0
			2610	1638	467	490	15			

- Molecule 2 is a protein called Envelope glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	150	Total	C	N	O	S	0	0	0
			1142	708	198	231	5			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	522	SER	-	linker	UNP G4WYG8
B	523	GLY	-	linker	UNP G4WYG8
B	524	ARG	-	linker	UNP G4WYG8
B	525	LEU	-	linker	UNP G4WYG8
B	526	VAL	-	linker	UNP G4WYG8
B	527	PRO	-	linker	UNP G4WYG8
B	528	ARG	-	linker	UNP G4WYG8
B	529	GLY	-	linker	UNP G4WYG8
B	530	SER	-	linker	UNP G4WYG8
B	531	PRO	-	linker	UNP G4WYG8
B	532	GLY	-	linker	UNP G4WYG8
B	533	SER	-	linker	UNP G4WYG8
B	562	HIS	-	expression tag	UNP M1E1E4
B	563	HIS	-	expression tag	UNP M1E1E4
B	564	HIS	-	expression tag	UNP M1E1E4
B	565	HIS	-	expression tag	UNP M1E1E4
B	566	HIS	-	expression tag	UNP M1E1E4
B	567	HIS	-	expression tag	UNP M1E1E4

- Molecule 3 is a protein called SD84h.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	116	Total	C	N	O	S	0	0	0
			883	554	151	174	4			

- Molecule 4 is a protein called CR9114 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	209	Total	C	N	O	S	0	0	0
			1553	969	263	317	4			

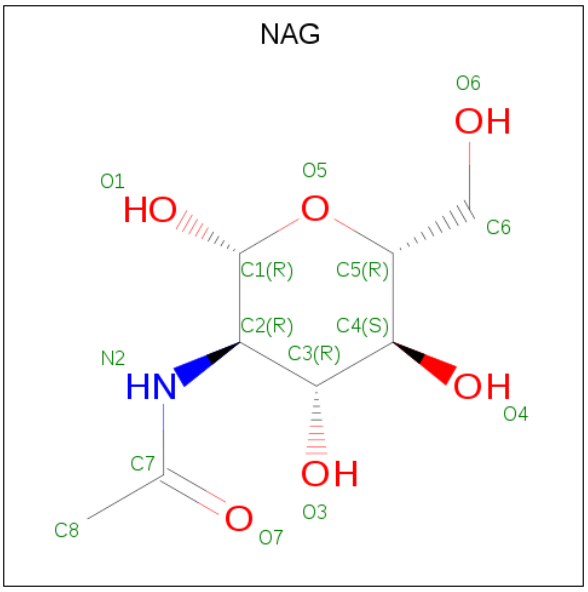
- Molecule 5 is a protein called CR9114 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	H	214	Total	C	N	O	S	0	0	0
			1587	1002	264	314	7			

- Molecule 6 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
6	D	2	Total	C	N	O	0	0	0
			28	16	2	10			
6	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
6	F	2	Total	C	N	O	0	0	0
			28	16	2	10			
6	G	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



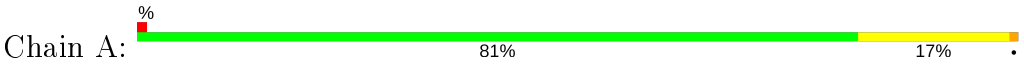
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	N	O	0	0
			14	8	1	5		

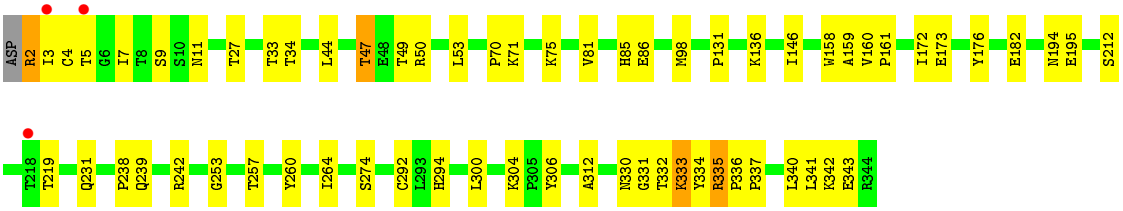
3 Residue-property plots (i)

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

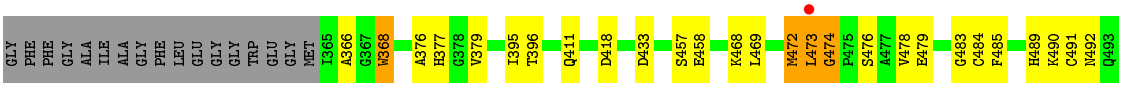
Chain A:

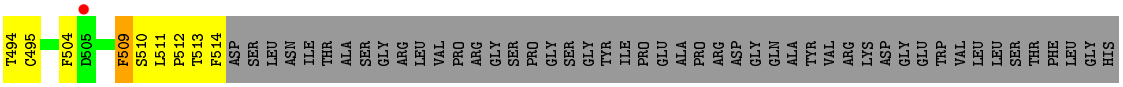




- Molecule 2: Envelope glycoprotein

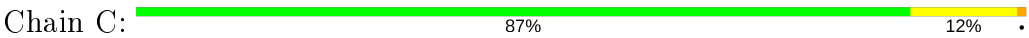
Chain B:

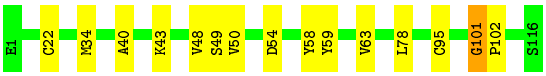




- Molecule 3: SD84h

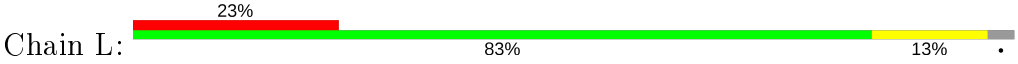
Chain C:

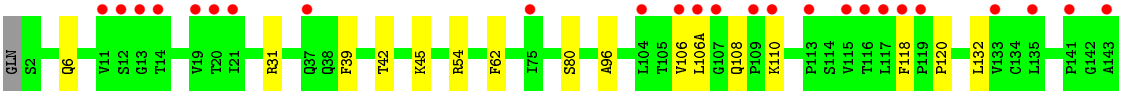


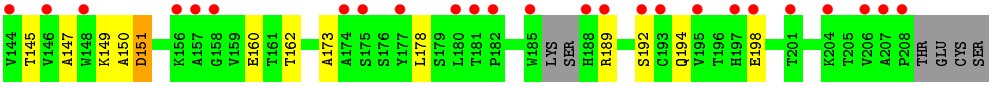


- Molecule 4: CR9114 Light chain

Chain L:

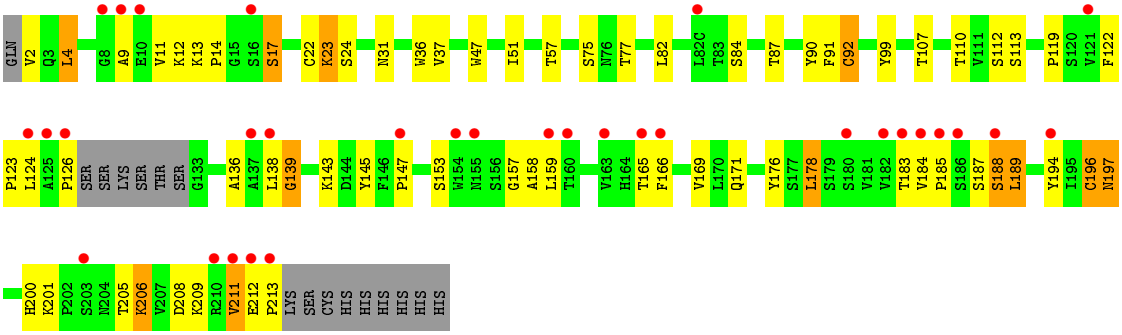






• Molecule 5: CR9114 Fab heavy chain

Chain H:



• Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

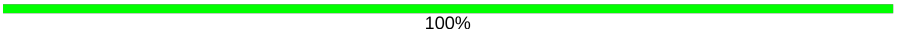
Chain D:

• Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:

• Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:



• Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α, β, γ	190.78Å 190.78Å 190.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.70 – 4.10 47.70 – 3.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.70-4.10) 100.0 (47.70-3.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.25 (at 3.77Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.240 , 0.276 0.243 , 0.277	Depositor DCC
R_{free} test set	1152 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	149.5	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 163.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.056 for l,-k,h	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	7901	wwPDB-VP
Average B, all atoms (Å^2)	188.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.17% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	0.23	0/2670	0.50	0/3629
2	B	0.24	0/1156	0.49	0/1559
3	C	0.22	0/903	0.46	0/1226
4	L	0.22	0/1590	0.52	0/2171
5	H	0.32	0/1626	0.86	7/2217 (0.3%)
All	All	0.25	0/7945	0.59	7/10802 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
5	H	0	1
All	All	0	2

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	H	178	LEU	CA-CB-CG	7.99	133.68	115.30
5	H	139	GLY	N-CA-C	7.09	130.84	113.10
5	H	4	LEU	CA-CB-CG	6.00	129.10	115.30
5	H	157	GLY	C-N-CA	-5.73	107.38	121.70
5	H	23	LYS	C-N-CA	-5.65	107.57	121.70
5	H	92	CYS	N-CA-C	-5.19	96.99	111.00
5	H	159	LEU	CB-CG-CD1	5.02	119.54	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	336	PRO	Peptide
5	H	158	ALA	Mainchain

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	2610	0	2597	37	0
2	B	1142	0	1125	19	0
3	C	883	0	849	9	0
4	L	1553	0	1498	16	0
5	H	1587	0	1537	42	0
6	D	28	0	25	0	0
6	E	28	0	25	0	0
6	F	28	0	25	0	0
6	G	28	0	25	1	0
7	A	14	0	13	0	0
All	All	7901	0	7719	105	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

All (105) 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)
5:H:126:PRO:HD2	5:H:213:PRO:HA	1.61	0.83
1:A:2:ARG:NH2	1:A:340:LEU:O	2.11	0.82
5:H:124:LEU:HD12	5:H:139:GLY:HA3	1.61	0.81
5:H:9:ALA:H	5:H:201:LYS:HD2	1.47	0.79
2:B:376:ALA:H	2:B:490:LYS:HE2	1.51	0.75
5:H:126:PRO:HG3	5:H:138:LEU:HB3	1.67	0.75
5:H:119:PRO:HB3	5:H:145:TYR:HB3	1.73	0.71
4:L:162:THR:HG22	5:H:169:VAL:HB	1.72	0.70
5:H:17:SER:HB3	5:H:82:LEU:O	1.94	0.68
1:A:194:ASN:OD1	1:A:195:GLU:N	2.28	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:238:PRO:HG2	3:C:50:VAL:HG11	1.77	0.67
5:H:185:PRO:HG2	5:H:188:SER:HB2	1.77	0.66
5:H:124:LEU:O	5:H:211:VAL:HG21	1.96	0.65
4:L:147:ALA:HB3	4:L:194:GLN:HB2	1.79	0.64
1:A:4:CYS:HA	2:B:484:CYS:HA	1.82	0.62
5:H:200:HIS:HB3	5:H:205:THR:HB	1.79	0.62
1:A:146:ILE:HG22	6:G:1:NAG:H82	1.82	0.62
5:H:90:TYR:N	5:H:107:THR:O	2.33	0.61
4:L:149:LYS:HB2	4:L:192:SER:HB2	1.83	0.61
5:H:153:SER:OG	5:H:197:ASN:OD1	2.21	0.59
4:L:80:SER:HA	4:L:106:VAL:HG11	1.85	0.59
4:L:132:LEU:HB2	4:L:178:LEU:HB3	1.84	0.59
2:B:396:THR:HG21	5:H:31:ASN:HB3	1.83	0.59
1:A:330:ASN:HA	2:B:395:ILE:HD13	1.85	0.58
1:A:158:TRP:CZ3	3:C:101:GLY:HA3	2.39	0.58
5:H:22:CYS:O	5:H:77:THR:HG23	2.07	0.55
3:C:40:ALA:HB3	3:C:43:LYS:HD2	1.89	0.55
2:B:509:PHE:O	2:B:511:LEU:N	2.40	0.54
5:H:13:LYS:HD2	5:H:113:SER:HA	1.90	0.54
4:L:54:ARG:HD3	4:L:62:PHE:O	2.07	0.53
2:B:476:SER:HB3	2:B:512:PRO:HD2	1.91	0.53
4:L:118:PHE:CD1	5:H:124:LEU:HB3	2.44	0.52
4:L:118:PHE:CG	5:H:124:LEU:HB3	2.44	0.52
1:A:49:THR:HG23	1:A:81:VAL:HG12	1.91	0.52
1:A:50:ARG:HD3	1:A:53:LEU:HD23	1.91	0.52
1:A:176:TYR:CG	1:A:253:GLY:HA2	2.46	0.52
1:A:331:GLY:HA2	2:B:368:TRP:CZ2	2.44	0.51
3:C:49:SER:HG	3:C:59:TYR:HD2	1.57	0.51
5:H:197:ASN:ND2	5:H:208:ASP:OD1	2.37	0.51
1:A:160:VAL:HG21	3:C:102:PRO:HB3	1.92	0.51
1:A:294:HIS:CE1	1:A:312:ALA:HB3	2.46	0.50
1:A:11:ASN:OD1	1:A:337:PRO:HB3	2.12	0.50
2:B:468:LYS:O	2:B:472:MET:HG3	2.13	0.49
1:A:3:ILE:HD11	2:B:491:CYS:HB2	1.95	0.49
5:H:12:LYS:HE2	5:H:17:SER:O	2.13	0.48
4:L:118:PHE:CD2	5:H:124:LEU:HD13	2.48	0.48
2:B:366:ALA:HB2	5:H:99:TYR:CZ	2.49	0.48
1:A:33:THR:HG22	1:A:306:TYR:HB2	1.96	0.48
5:H:119:PRO:CB	5:H:145:TYR:HB3	2.42	0.48
5:H:14:PRO:HD3	5:H:112:SER:O	2.13	0.48
5:H:136:ALA:O	5:H:183:THR:HG23	2.13	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:212:SER:OG	1:A:219:THR:OG1	2.32	0.47
5:H:187:SER:C	5:H:189:LEU:H	2.18	0.47
2:B:473:LEU:HA	2:B:504:PHE:HE2	1.78	0.47
1:A:131:PRO:HG2	1:A:172:ILE:HG22	1.95	0.47
1:A:173:GLU:HG2	1:A:257:THR:HG22	1.97	0.47
5:H:122:PHE:CD2	5:H:143:LYS:HD3	2.50	0.47
3:C:50:VAL:HG13	3:C:58:TYR:HD1	1.80	0.47
1:A:27:THR:HG23	1:A:333:LYS:O	2.15	0.46
2:B:489:HIS:ND1	2:B:490:LYS:O	2.32	0.45
5:H:51:ILE:HG13	5:H:57:THR:HG22	1.97	0.45
5:H:122:PHE:HD2	5:H:143:LYS:HD3	1.82	0.45
1:A:5:THR:N	2:B:483:GLY:O	2.45	0.45
1:A:182:GLU:HG2	1:A:274:SER:HB3	1.99	0.45
2:B:377:HIS:CE1	2:B:492:ASN:H	2.35	0.45
4:L:39:PHE:O	4:L:42:THR:HG22	2.17	0.44
5:H:206:LYS:HD2	5:H:206:LYS:HA	1.63	0.44
1:A:9:SER:HB2	1:A:27:THR:HG21	1.98	0.44
1:A:44:LEU:O	1:A:47:THR:HG22	2.17	0.44
5:H:36:TRP:CZ3	5:H:92:CYS:HB3	2.53	0.44
4:L:96:ALA:HB3	5:H:47:TRP:CG	2.53	0.44
5:H:212:GLU:HA	5:H:213:PRO:HD3	1.84	0.44
5:H:23:LYS:O	5:H:23:LYS:HG3	2.18	0.44
1:A:176:TYR:CD1	1:A:253:GLY:HA2	2.53	0.43
3:C:48:VAL:HG13	3:C:63:VAL:HG11	2.01	0.43
5:H:4:LEU:HD23	5:H:24:SER:HB3	2.01	0.43
1:A:34:THR:HG23	1:A:304:LYS:HD2	2.01	0.43
5:H:87:THR:HG23	5:H:110:THR:HA	2.01	0.43
4:L:173:ALA:HB1	5:H:166:PHE:CE2	2.53	0.43
5:H:184:VAL:HG21	5:H:194:TYR:CZ	2.54	0.42
5:H:37:VAL:O	5:H:91:PHE:HB2	2.19	0.42
2:B:491:CYS:O	2:B:495:CYS:HB3	2.19	0.42
5:H:124:LEU:N	5:H:139:GLY:O	2.37	0.42
3:C:34:MET:HB3	3:C:78:LEU:HD22	2.02	0.42
1:A:159:ALA:HB3	1:A:264:ILE:HG23	2.01	0.42
1:A:172:ILE:HG23	1:A:260:TYR:HE2	1.85	0.42
4:L:150:ALA:O	4:L:151:ASP:HB2	2.18	0.42
1:A:333:LYS:HB2	2:B:458:GLU:OE2	2.19	0.42
1:A:85:HIS:ND1	1:A:86:GLU:HG3	2.35	0.41
4:L:110:LYS:HE3	4:L:198:GLU:HG3	2.02	0.41
2:B:472:MET:O	2:B:474:GLY:N	2.53	0.41
1:A:292:CYS:HB3	1:A:300:LEU:HD12	2.03	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:479:GLU:HB2	2:B:485:PHE:CE2	2.56	0.41
1:A:75:LYS:HA	1:A:75:LYS:HD3	1.84	0.41
4:L:120:PRO:HD3	4:L:132:LEU:HG	2.02	0.41
1:A:231:GLN:HG2	1:A:242:ARG:NH2	2.35	0.41
1:A:7:ILE:HG21	1:A:332:THR:HG21	2.02	0.41
5:H:196:CYS:N	5:H:209:LYS:O	2.23	0.41
4:L:160:GLU:OE1	5:H:171:GLN:HG2	2.21	0.41
1:A:70:PRO:O	1:A:71:LYS:HB2	2.20	0.41
3:C:22:CYS:HB3	3:C:78:LEU:HB3	2.03	0.41
1:A:136:LYS:HE2	1:A:161:PRO:O	2.21	0.40
5:H:184:VAL:HB	5:H:185:PRO:HD2	2.03	0.40
1:A:2:ARG:HB3	1:A:2:ARG:HE	1.38	0.40
2:B:474:GLY:HA2	2:B:514:PHE:CE2	2.57	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	344/347 (99%)	329 (96%)	9 (3%)	6 (2%)	9 42
2	B	148/220 (67%)	136 (92%)	9 (6%)	3 (2%)	7 39
3	C	114/116 (98%)	110 (96%)	3 (3%)	1 (1%)	17 54
4	L	205/216 (95%)	199 (97%)	5 (2%)	1 (0%)	29 67
5	H	210/230 (91%)	188 (90%)	16 (8%)	6 (3%)	4 32
All	All	1021/1129 (90%)	962 (94%)	42 (4%)	17 (2%)	9 42

All (17) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	473	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	510	SER
3	C	101	GLY
5	H	84	SER
1	A	333	LYS
2	B	474	GLY
1	A	341	LEU
5	H	123	PRO
5	H	188	SER
5	H	206	LYS
1	A	334	TYR
4	L	151	ASP
5	H	147	PRO
1	A	342	LYS
1	A	343	GLU
5	H	165	THR
1	A	335	ARG

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	287/294 (98%)	282 (98%)	5 (2%)	60 78
2	B	124/177 (70%)	112 (90%)	12 (10%)	8 29
3	C	94/94 (100%)	92 (98%)	2 (2%)	53 72
4	L	173/180 (96%)	166 (96%)	7 (4%)	31 57
5	H	177/193 (92%)	167 (94%)	10 (6%)	21 49
All	All	855/938 (91%)	819 (96%)	36 (4%)	30 56

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	ARG
1	A	47	THR
1	A	98	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	239	GLN
1	A	335	ARG
2	B	368	TRP
2	B	379	VAL
2	B	411	GLN
2	B	418	ASP
2	B	433	ASP
2	B	457	SER
2	B	469	LEU
2	B	472	MET
2	B	478	VAL
2	B	494	THR
2	B	509	PHE
2	B	513	THR
3	C	54	ASP
3	C	95	CYS
4	L	6	GLN
4	L	31	ARG
4	L	45	LYS
4	L	106(A)	LEU
4	L	108	GLN
4	L	145	THR
4	L	189	ARG
5	H	2	VAL
5	H	11	VAL
5	H	17	SER
5	H	75	SER
5	H	176	TYR
5	H	178	LEU
5	H	189	LEU
5	H	196	CYS
5	H	197	ASN
5	H	211	VAL

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

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (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	$\#\|Z\| > 2$
6	NAG	D	1	1,6	14,14,15	0.34	0	17,19,21	0.37	0
6	NAG	D	2	6	14,14,15	0.65	1 (7%)	17,19,21	0.60	0
6	NAG	E	1	1,6	14,14,15	0.27	0	17,19,21	0.54	0
6	NAG	E	2	6	14,14,15	1.02	1 (7%)	17,19,21	0.84	1 (5%)
6	NAG	F	1	1,6	14,14,15	0.36	0	17,19,21	0.41	0
6	NAG	F	2	6	14,14,15	0.22	0	17,19,21	0.38	0
6	NAG	G	1	1,6	14,14,15	0.19	0	17,19,21	0.45	0
6	NAG	G	2	6	14,14,15	0.26	0	17,19,21	0.35	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	D	1	1,6	-	0/6/23/26	0/1/1/1
6	NAG	D	2	6	-	2/6/23/26	0/1/1/1
6	NAG	E	1	1,6	-	1/6/23/26	0/1/1/1
6	NAG	E	2	6	-	2/6/23/26	0/1/1/1
6	NAG	F	1	1,6	-	2/6/23/26	0/1/1/1
6	NAG	F	2	6	-	2/6/23/26	0/1/1/1
6	NAG	G	1	1,6	-	0/6/23/26	0/1/1/1
6	NAG	G	2	6	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	E	2	NAG	O5-C1	-3.61	1.38	1.43
6	D	2	NAG	O5-C1	-2.24	1.40	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	E	2	NAG	C3-C4-C5	2.57	114.83	110.24

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	E	2	NAG	O5-C5-C6-O6
6	D	2	NAG	O5-C5-C6-O6
6	E	2	NAG	C4-C5-C6-O6
6	F	2	NAG	C4-C5-C6-O6
6	F	2	NAG	O5-C5-C6-O6
6	G	2	NAG	C4-C5-C6-O6
6	D	2	NAG	C4-C5-C6-O6
6	F	1	NAG	O5-C5-C6-O6
6	G	2	NAG	O5-C5-C6-O6
6	F	1	NAG	C4-C5-C6-O6
6	E	1	NAG	C4-C5-C6-O6

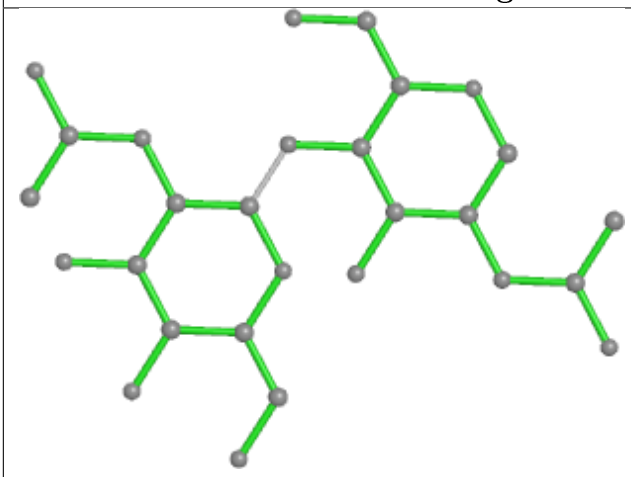
There are no ring outliers.

1 monomer is involved in 1 short contact:

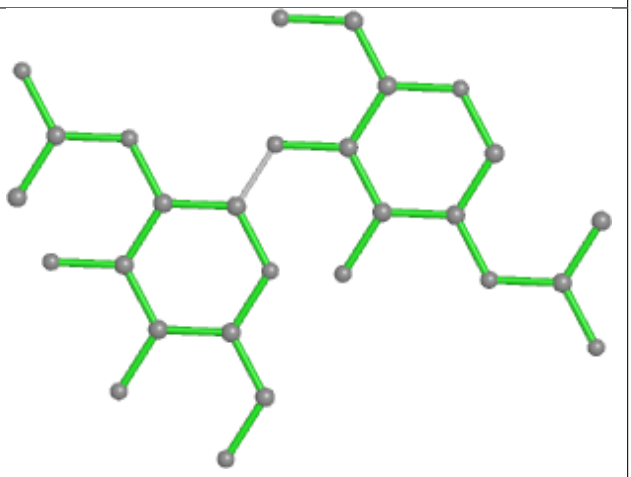
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	G	1	NAG	1	0

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

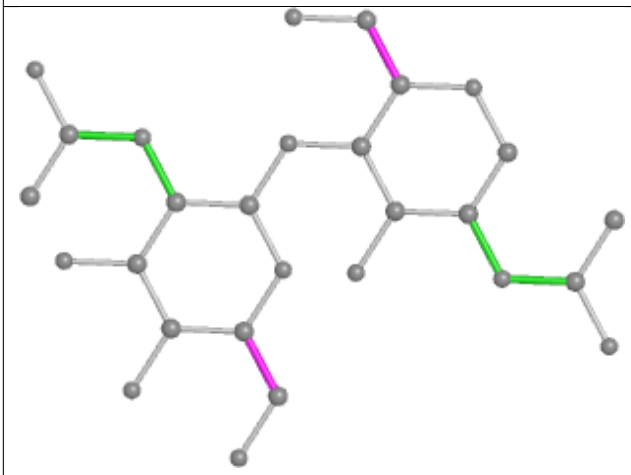
Oligosaccharide Chain F



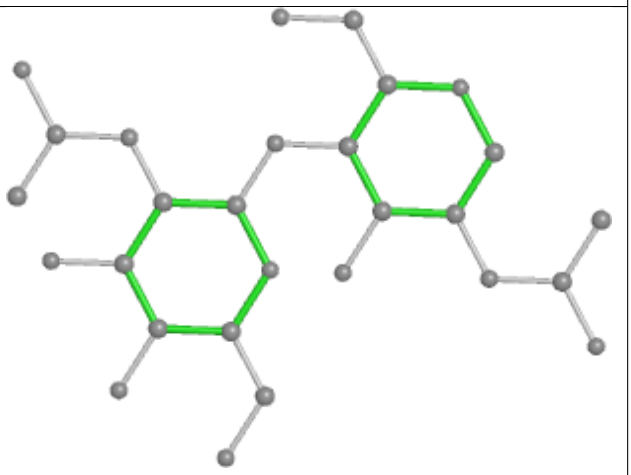
Bond lengths



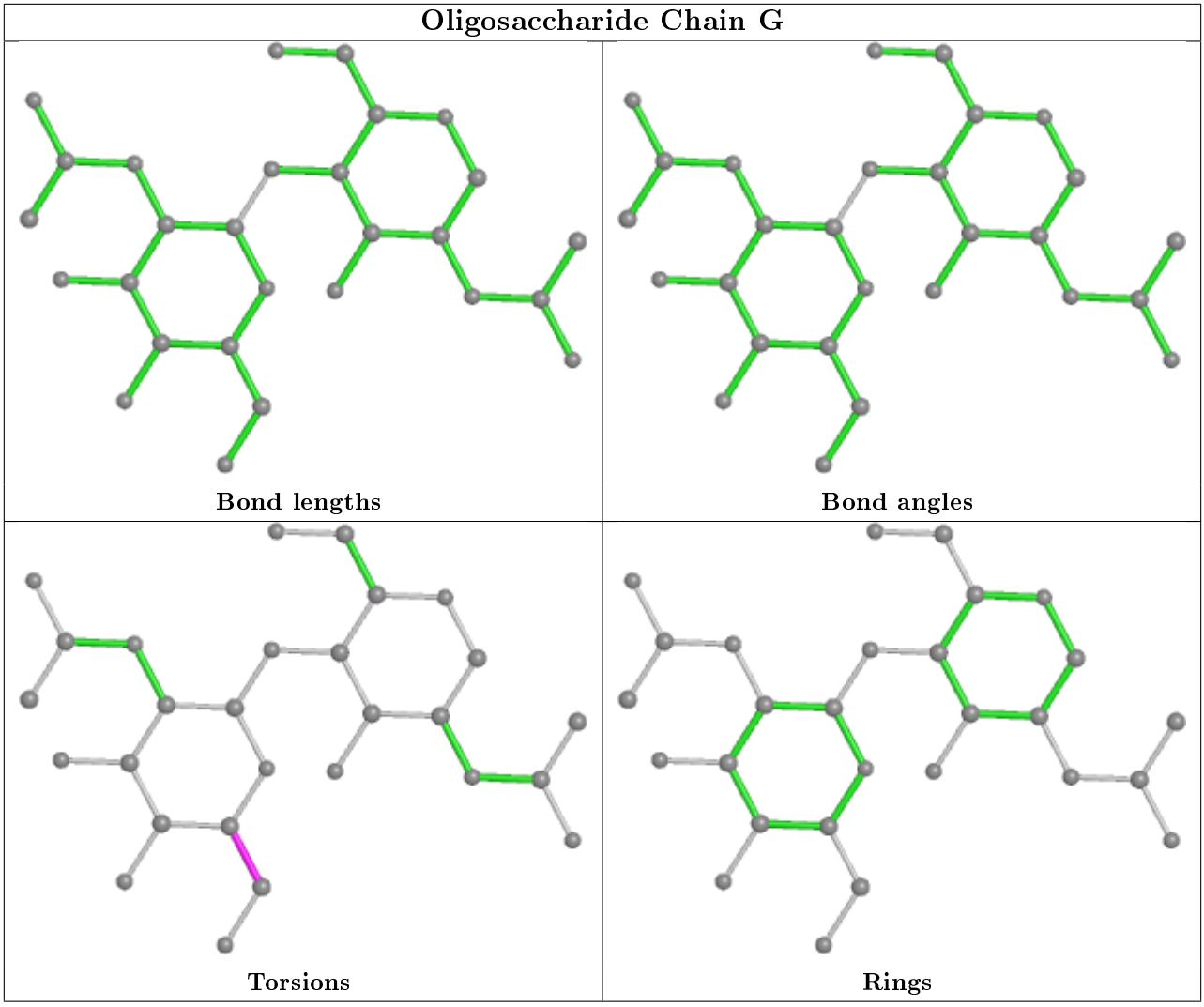
Bond angles



Torsions



Rings



5.6 Ligand geometry (i)

1 ligand is 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	$\#\|Z\| > 2$
7	NAG	A	409	1	14,14,15	0.27	0	17,19,21	0.47	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	NAG	A	409	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 95^{th} 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	346/347 (99%)	-0.01	3 (0%) 84 77	97, 134, 189, 232	0
2	B	150/220 (68%)	0.23	2 (1%) 77 68	94, 176, 240, 256	0
3	C	116/116 (100%)	0.18	0 100 100	126, 156, 189, 216	0
4	L	209/216 (96%)	1.00	50 (23%) 0 1	164, 263, 322, 336	0
5	H	214/230 (93%)	0.97	32 (14%) 2 3	132, 232, 320, 335	0
All	All	1035/1129 (91%)	0.45	87 (8%) 11 10	94, 171, 310, 336	0

All (87) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	H	212	GLU	12.3
5	H	213	PRO	11.2
5	H	211	VAL	10.2
5	H	126	PRO	10.0
5	H	184	VAL	8.3
4	L	144	VAL	7.2
4	L	182	PRO	6.9
4	L	118	PHE	5.9
4	L	193	CYS	5.4
4	L	117	LEU	5.3
5	H	183	THR	5.1
4	L	180	LEU	5.0
5	H	194	TYR	4.9
4	L	207	ALA	4.9
4	L	12	SER	4.8
4	L	206	VAL	4.4
5	H	138	LEU	4.2
4	L	189	ARG	3.8
5	H	10	GLU	3.8
4	L	197	HIS	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	L	106(A)	LEU	3.6
5	H	180	SER	3.6
4	L	133	VAL	3.4
5	H	137	ALA	3.4
4	L	14	THR	3.3
5	H	186	SER	3.2
5	H	185	PRO	3.2
5	H	9	ALA	3.2
4	L	188	HIS	3.2
5	H	188	SER	3.1
4	L	208	PRO	3.1
4	L	37	GLN	3.0
5	H	160	THR	3.0
5	H	165	THR	3.0
5	H	210	ARG	2.9
4	L	156	LYS	2.9
5	H	8	GLY	2.9
5	H	182	VAL	2.9
4	L	146	VAL	2.9
2	B	473	LEU	2.8
4	L	113	PRO	2.8
4	L	157	ALA	2.7
4	L	198	GLU	2.7
4	L	106	VAL	2.7
4	L	143	ALA	2.7
4	L	185	TRP	2.6
5	H	159	LEU	2.6
4	L	195	VAL	2.6
5	H	163	VAL	2.6
5	H	121	VAL	2.6
4	L	177	TYR	2.6
4	L	119	PRO	2.5
4	L	21	ILE	2.5
1	A	218	THR	2.5
4	L	141	PRO	2.5
5	H	125	ALA	2.5
4	L	148	TRP	2.5
2	B	505	ASP	2.4
4	L	181	THR	2.4
4	L	175	SER	2.3
5	H	82(C)	LEU	2.3
4	L	158	GLY	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
5	H	124	LEU	2.3
1	A	3	ILE	2.3
4	L	174	ALA	2.3
4	L	204	LYS	2.2
1	A	5	THR	2.2
4	L	115	VAL	2.2
5	H	166	PHE	2.2
5	H	16	SER	2.2
4	L	11	VAL	2.2
4	L	13	GLY	2.1
4	L	116	THR	2.1
4	L	192	SER	2.1
4	L	75	ILE	2.1
4	L	20	THR	2.1
5	H	147	PRO	2.1
4	L	19	VAL	2.1
4	L	107	GLY	2.1
4	L	104	LEU	2.1
5	H	155	ASN	2.1
4	L	110	LYS	2.0
5	H	203	SER	2.0
4	L	109	PRO	2.0
4	L	201	THR	2.0
5	H	154	TRP	2.0
4	L	135	LEU	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, 95^{th} 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	F	2	14/15	0.64	0.46	209,270,283,284	0
6	NAG	G	2	14/15	0.68	0.36	198,233,242,244	0
6	NAG	G	1	14/15	0.79	0.36	189,211,232,253	0

Continued on next page...

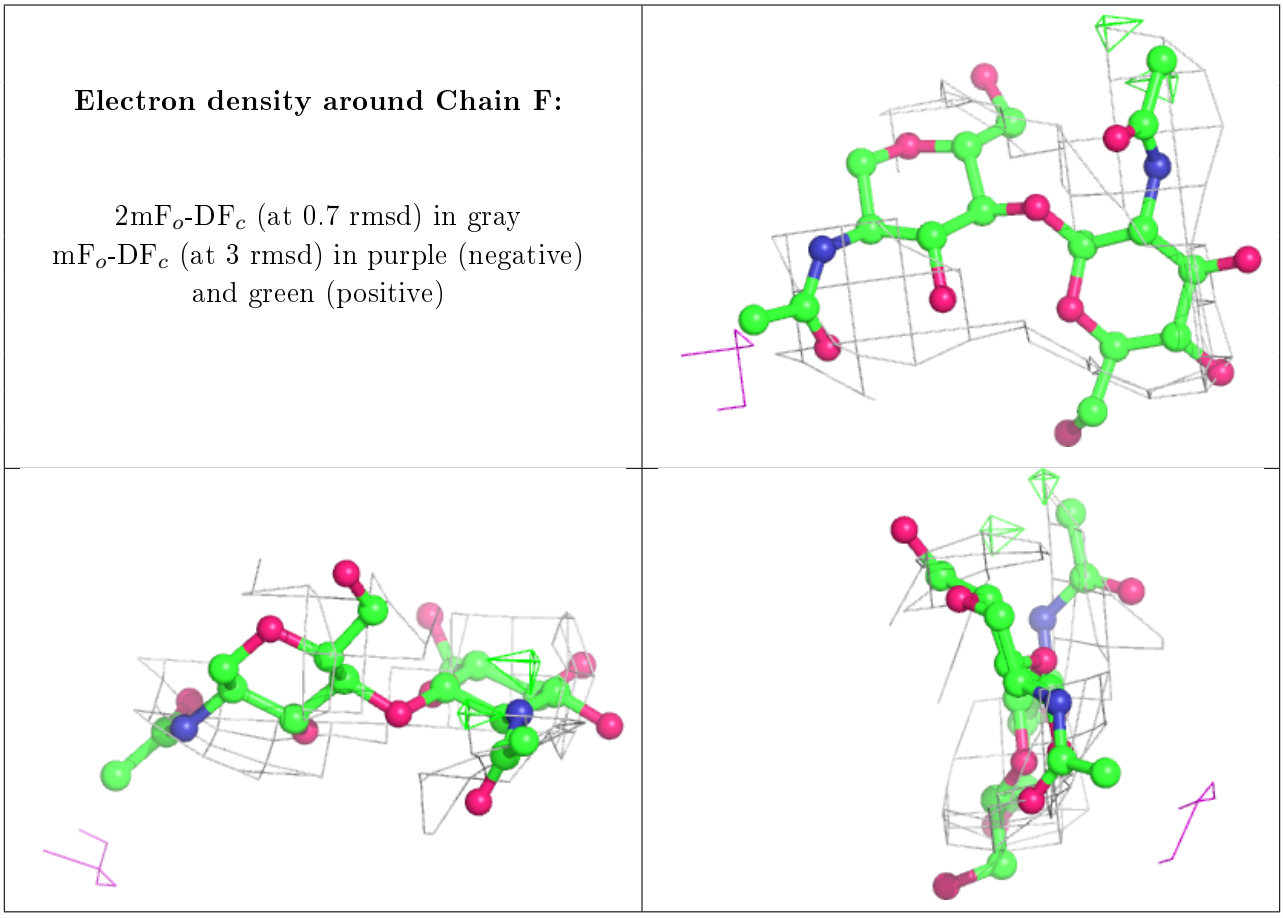
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	NAG	E	2	14/15	0.86	0.23	183,197,242,242	0
6	NAG	E	1	14/15	0.86	0.28	123,155,201,207	0
6	NAG	D	1	14/15	0.87	0.20	137,170,188,195	0
6	NAG	F	1	14/15	0.87	0.19	230,242,256,271	0
6	NAG	D	2	14/15	0.89	0.20	163,202,214,217	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.

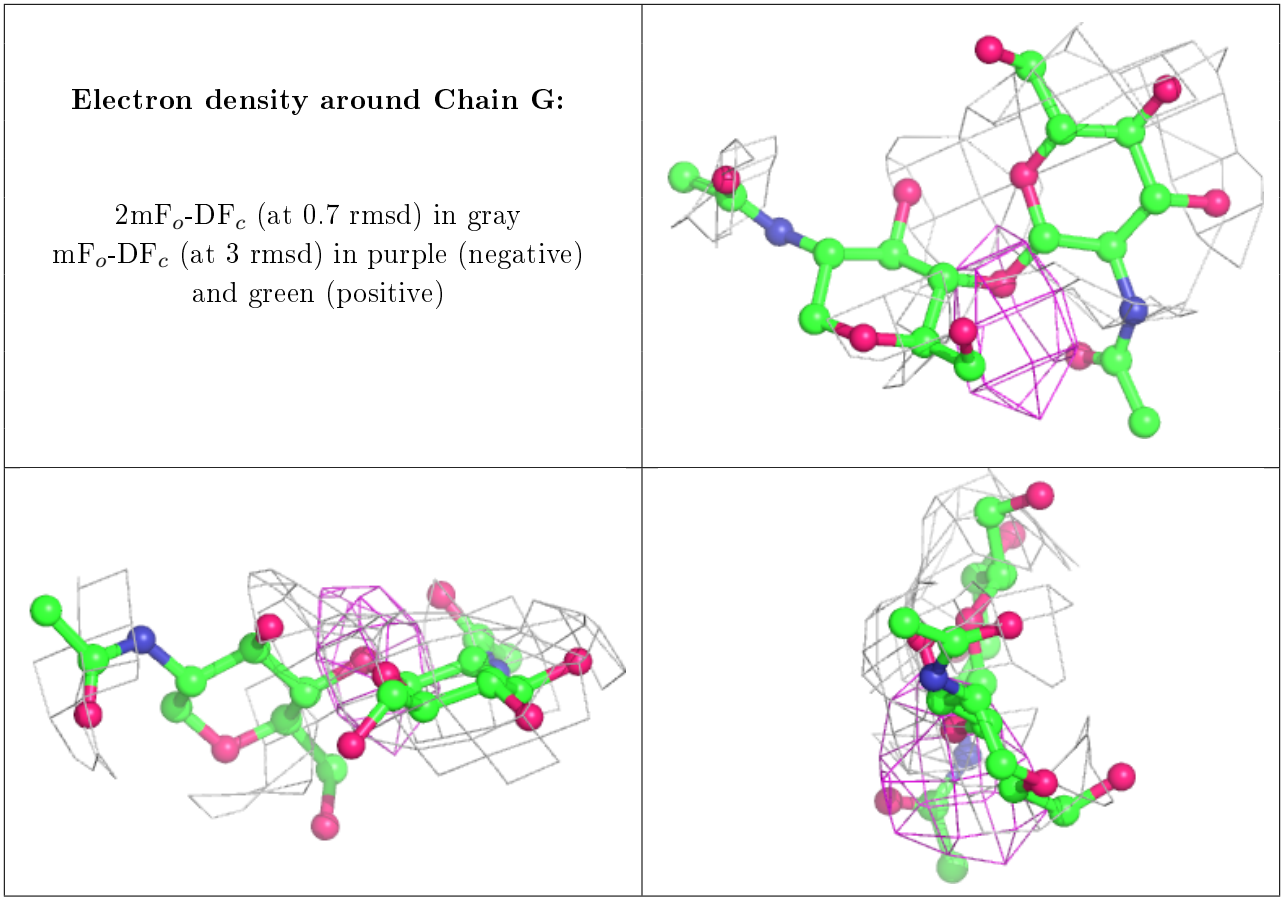
Electron density around Chain F:

$2mF_o$-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 G:

$2mF_o$-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, 95^{th} 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
7	NAG	A	409	14/15	0.76	0.34	163,200,221,229	0

6.5 Other polymers

There are no such residues in this entry.