



Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 04:48 PM BST

PDB ID : 4N5K
Title : Crystal structure of hemagglutinin from an H7N9 influenza virus in complex with LSTa
Authors : Xu, R.; Wilson, I.A.
Deposited on : 2013-10-10
Resolution : 2.71 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:

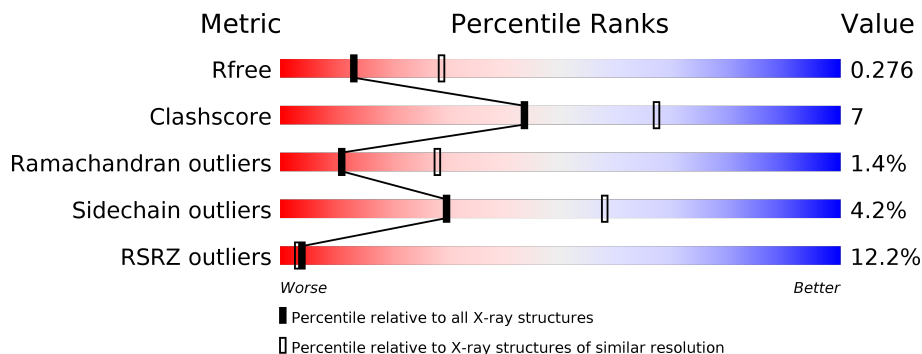
MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.13.1
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & 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 2.71 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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

Mol	Chain	Length	Quality of chain
1	A	321	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 16%, green 81%, yellow 97%, orange 98%, red 99%);"></div> <div style="margin-left: 10px;"> <p>16% 81% 16% ...</p> </div> </div>
1	C	321	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 16%, orange 16%, yellow 76%, green 76%, yellow 92%, orange 95%, red 98%);"></div> <div style="margin-left: 10px;"> <p>16% 76% 22% ..</p> </div> </div>
2	B	180	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 84%, green 84%, yellow 92%, orange 95%, red 98%, grey 100%);"></div> <div style="margin-left: 10px;"> <p>84% 8% 6%</p> </div> </div>
2	D	180	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 35%, orange 35%, yellow 73%, green 73%, yellow 81%, orange 84%, red 87%, grey 100%);"></div> <div style="margin-left: 10px;"> <p>35% 73% 18% .. 6%</p> </div> </div>
3	E	3	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 33%, yellow 67%);"></div> <div style="margin-left: 10px;"> <p>33% 67%</p> </div> </div>
4	F	2	<div style="display: flex; align-items: center;"> <div style="width: 100%; height: 10px; background: linear-gradient(to right, orange 100%);"></div> <div style="margin-left: 10px;"> <p>100%</p> </div> </div>

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 7760 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	318	2427	1509	438	465	15	0	0	0
1	C	318	2427	1509	438	465	15	0	0	0

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	169	1373	845	241	280	7	0	0	0
2	D	169	1373	845	241	280	7	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	177	SER	-	expression tag	UNP R4NN21
B	178	GLY	-	expression tag	UNP R4NN21
B	179	ARG	-	expression tag	UNP R4NN21
B	180	LEU	-	expression tag	UNP R4NN21
B	181	VAL	-	expression tag	UNP R4NN21
B	182	PRO	-	expression tag	UNP R4NN21
B	183	ARG	-	expression tag	UNP R4NN21
D	177	SER	-	expression tag	UNP R4NN21
D	178	GLY	-	expression tag	UNP R4NN21
D	179	ARG	-	expression tag	UNP R4NN21
D	180	LEU	-	expression tag	UNP R4NN21
D	181	VAL	-	expression tag	UNP R4NN21
D	182	PRO	-	expression tag	UNP R4NN21
D	183	ARG	-	expression tag	UNP R4NN21

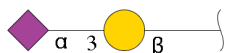
- Molecule 3 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-b

eta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



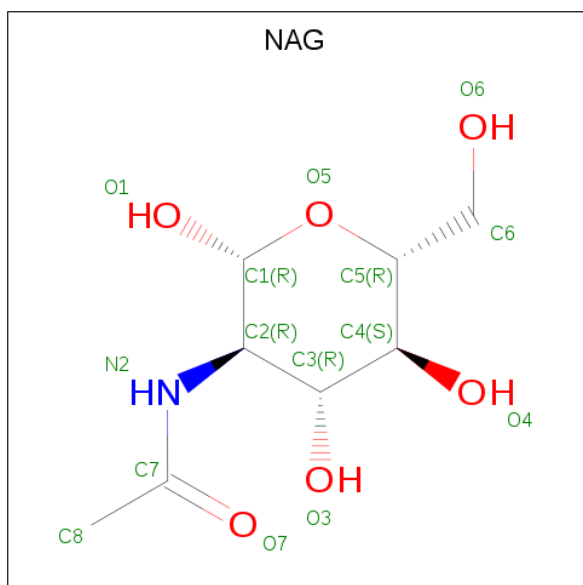
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	E	3	39	22	2	15	0	0	0

- Molecule 4 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	F	2	31	17	1	13	0	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



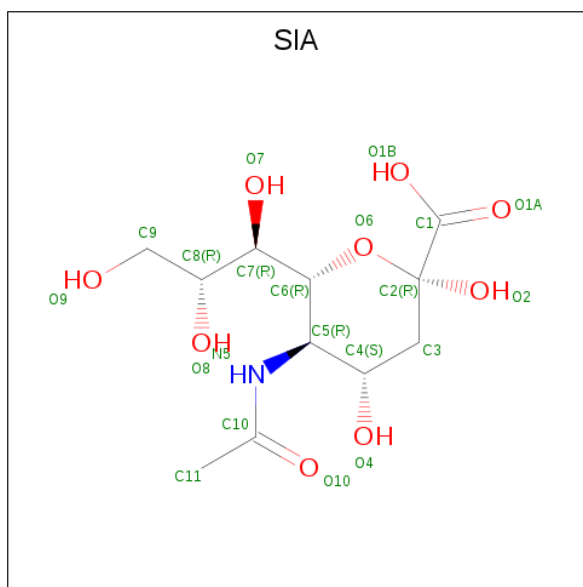
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	A	1	14	8	1	5	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 6 is N-acetyl-alpha-neuraminic acid (three-letter code: SIA) (formula: C₁₁H₁₉NO₉).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	C	1	Total	C	N	O	0	0
			20	11	1	8		

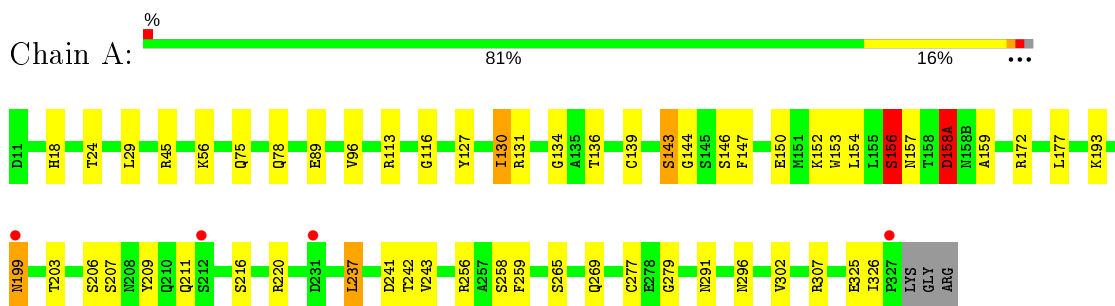
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	15	Total	O	0	0
			15	15		
7	B	12	Total	O	0	0
			12	12		
7	C	1	Total	O	0	0
			1	1		

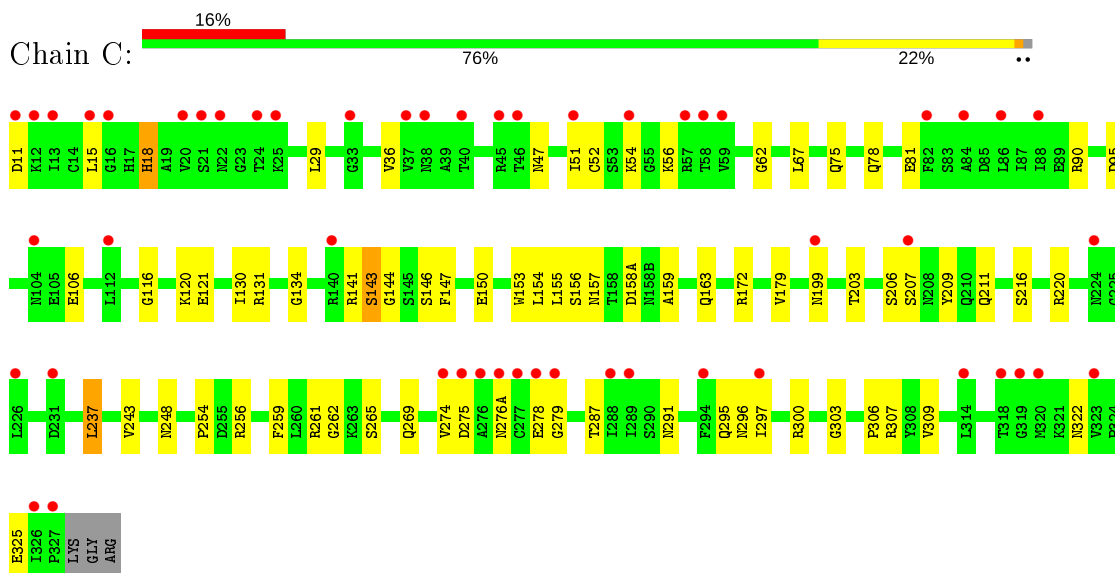
3 Residue-property plots [i](#)

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.

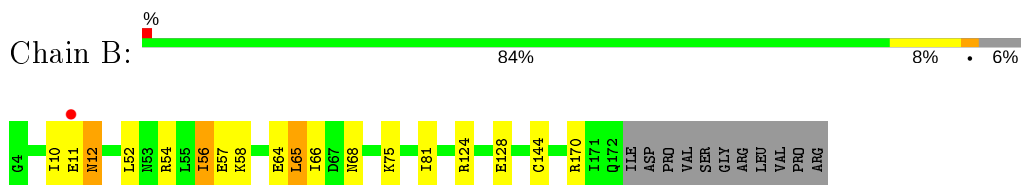
- Molecule 1: Hemagglutinin HA1



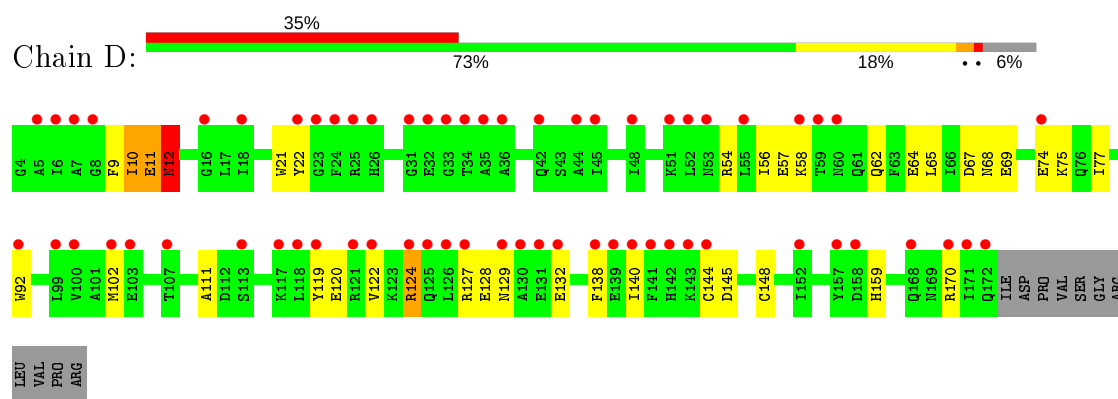
- Molecule 1: Hemagglutinin HA1



- Molecule 2: Hemagglutinin HA2



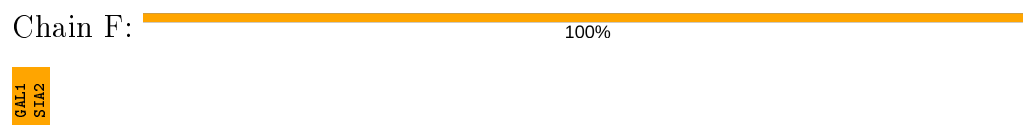
- Molecule 2: Hemagglutinin HA2



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



- Molecule 4: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α , β , γ	154.70Å 154.70Å 154.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.92 – 2.71 48.92 – 2.71	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.92-2.71) 99.5 (48.92-2.71)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.13 (at 2.69Å)	Xtrriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.227 , 0.276 0.227 , 0.276	Depositor DCC
R_{free} test set	1718 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	68.4	Xtrriage
Anisotropy	0.000	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 55.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.032 for l,-k,h	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7760	wwPDB-VP
Average B, all atoms (Å ²)	100.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.96% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SIA, GAL, BMA, 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	0.26	0/2474	0.43	0/3345
1	C	0.23	0/2474	0.40	0/3345
2	B	0.27	0/1396	0.40	0/1881
2	D	0.23	0/1396	0.37	0/1881
All	All	0.25	0/7740	0.40	0/10452

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [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	2427	0	2387	34	0
1	C	2427	0	2389	49	0
2	B	1373	0	1273	12	0
2	D	1373	0	1273	28	0
3	E	39	0	34	0	0
4	F	31	0	26	1	0
5	A	14	0	13	0	0
5	B	14	0	13	1	0
5	D	14	0	13	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	C	20	0	17	1	0
7	A	15	0	0	0	0
7	B	12	0	0	0	0
7	C	1	0	0	0	0
All	All	7760	0	7438	102	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 (102) 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)
2:B:128:GLU:HG3	2:B:170:ARG:HH12	1.39	0.84
1:C:291:ASN:OD1	2:D:58:LYS:NZ	2.11	0.83
1:A:158(A):ASP:OD2	1:A:193:LYS:NZ	2.12	0.81
1:C:300:ARG:NH1	2:D:69:GLU:OE2	2.11	0.80
2:B:128:GLU:HG3	2:B:170:ARG:NH1	2.01	0.76
1:A:56:LYS:NZ	1:A:279:GLY:O	2.22	0.73
1:A:216:SER:O	1:A:220:ARG:NH2	2.19	0.70
1:A:326:ILE:HG12	2:B:12:ASN:HB3	1.73	0.70
1:C:216:SER:O	1:C:220:ARG:NH2	2.25	0.68
1:C:296:ASN:ND2	1:C:309:VAL:O	2.25	0.68
1:A:291:ASN:OD1	2:B:58:LYS:NZ	2.26	0.67
1:C:150:GLU:OE1	1:C:256:ARG:NH1	2.26	0.67
2:D:11:GLU:O	2:D:12:ASN:ND2	2.28	0.65
1:A:45:ARG:NH1	1:A:296:ASN:OD1	2.30	0.65
1:A:177:LEU:HB3	1:A:258:SER:HB2	1.78	0.64
1:A:307:ARG:HG2	1:A:307:ARG:HH11	1.62	0.63
1:A:134:GLY:HA3	1:A:153:TRP:HB3	1.79	0.62
1:C:303:GLY:HA2	2:D:62:GLN:HG3	1.81	0.62
1:A:325:GLU:HA	2:B:12:ASN:HB2	1.82	0.61
2:D:119:TYR:OH	2:D:132:GLU:OE2	2.12	0.61
2:B:128:GLU:O	2:B:170:ARG:NH1	2.33	0.61
1:C:11:ASP:OD1	2:D:144:CYS:N	2.32	0.60
1:C:47:ASN:HB2	1:C:297:ILE:HD13	1.84	0.59
2:D:145:ASP:OD2	2:D:148:CYS:N	2.28	0.58
1:A:237:LEU:HG	1:A:243:VAL:HG13	1.85	0.58
1:C:81:GLU:O	1:C:261:ARG:NH2	2.35	0.57
1:A:206:SER:OG	1:A:207:SER:N	2.36	0.56
1:C:15:LEU:HD21	2:D:122:VAL:HG21	1.88	0.56
1:A:172:ARG:HD3	1:A:259:PHE:CZ	2.42	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:54:LYS:HB3	1:C:278:GLU:HB2	1.88	0.55
1:C:206:SER:OG	1:C:207:SER:N	2.34	0.54
1:A:206:SER:HB3	1:A:209:TYR:HB3	1.89	0.54
1:C:325:GLU:HA	2:D:12:ASN:HB2	1.88	0.54
1:C:121:GLU:OE2	1:C:172:ARG:HD2	2.07	0.54
1:A:136:THR:HG23	1:A:139:CYS:H	1.72	0.53
2:D:120:GLU:HB3	2:D:124:ARG:NH1	2.23	0.53
4:F:1:GAL:O2	4:F:2:SIA:O1B	2.26	0.53
2:B:66:ILE:HG23	2:B:81:ILE:HD13	1.90	0.53
1:A:130:ILE:HD11	1:A:154:LEU:HB3	1.90	0.53
1:C:134:GLY:HA3	1:C:153:TRP:HB3	1.90	0.52
1:C:237:LEU:HG	1:C:243:VAL:HG13	1.92	0.51
1:C:146:SER:OG	1:C:147:PHE:N	2.43	0.51
2:B:75:LYS:HD3	5:B:201:NAG:H81	1.93	0.51
1:C:56:LYS:NZ	1:C:279:GLY:O	2.43	0.51
1:C:206:SER:HB3	1:C:209:TYR:HB3	1.92	0.50
1:A:291:ASN:HA	2:B:58:LYS:NZ	2.27	0.49
1:A:199:ASN:OD1	1:A:199:ASN:N	2.43	0.49
1:A:156:SER:OG	1:A:159:ALA:O	2.23	0.48
1:C:131:ARG:HB2	1:C:155:LEU:HB2	1.95	0.48
1:C:52:CYS:HB2	1:C:287:THR:HG21	1.95	0.48
1:C:275:ASP:OD2	1:C:276(A):ASN:N	2.40	0.48
1:C:291:ASN:HA	2:D:58:LYS:NZ	2.29	0.48
2:D:9:PHE:CD1	2:D:10:ILE:HG13	2.49	0.47
1:C:29:LEU:HD21	2:D:102:MET:HA	1.96	0.47
1:A:116:GLY:HA2	1:A:265:SER:HB3	1.97	0.47
1:C:172:ARG:HD3	1:C:259:PHE:CZ	2.50	0.47
2:D:75:LYS:HD3	5:D:201:NAG:H81	1.96	0.47
1:A:291:ASN:HB3	2:B:56:ILE:HG13	1.97	0.47
1:C:120:LYS:NZ	1:C:150:GLU:OE2	2.26	0.46
1:A:143:SER:OG	1:A:144:GLY:N	2.48	0.46
1:C:295:GLN:HG2	1:C:306:PRO:HG2	1.97	0.46
1:A:127:TYR:HB3	1:A:130:ILE:HG21	1.97	0.46
2:D:128:GLU:HG3	2:D:170:ARG:NH1	2.31	0.46
1:A:146:SER:OG	1:A:147:PHE:N	2.45	0.46
1:C:199:ASN:N	1:C:199:ASN:OD1	2.47	0.46
1:C:291:ASN:HB3	2:D:56:ILE:HG13	1.98	0.45
1:A:241:ASP:OD1	1:A:242:THR:N	2.47	0.45
1:C:18:HIS:N	2:D:21:TRP:O	2.37	0.45
1:C:106:GLU:OE2	2:D:68:ASN:HB2	2.17	0.45
1:A:89:GLU:OE2	1:A:113:ARG:NE	2.42	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:269:GLN:NE2	2:D:68:ASN:OD1	2.39	0.44
1:C:116:GLY:HA2	1:C:265:SER:HB3	2.00	0.44
1:C:307:ARG:HG3	2:D:92:TRP:CZ3	2.53	0.44
1:C:163:GLN:OE1	1:C:248:ASN:HB3	2.17	0.44
1:C:150:GLU:OE1	1:C:256:ARG:HD3	2.18	0.43
1:A:130:ILE:HA	1:A:156:SER:O	2.17	0.43
1:C:36:VAL:HA	1:C:322:ASN:H	1.82	0.43
1:A:150:GLU:OE1	1:A:256:ARG:HD3	2.18	0.43
1:C:62:GLY:HA2	1:C:90:ARG:HG3	1.99	0.43
6:C:401:SIA:O8	6:C:401:SIA:O6	2.33	0.43
1:C:51:ILE:HB	1:C:274:VAL:HA	2.02	0.42
1:C:18:HIS:CG	2:D:21:TRP:HA	2.55	0.42
2:D:22:TYR:OH	2:D:111:ALA:O	2.35	0.42
1:A:152:LYS:HA	1:A:152:LYS:HD3	1.90	0.42
1:A:130:ILE:HG12	1:A:131:ARG:N	2.34	0.42
1:A:269:GLN:HE22	2:B:68:ASN:CG	2.22	0.41
1:C:143:SER:OG	1:C:144:GLY:N	2.52	0.41
2:D:132:GLU:HG2	2:D:138:PHE:CE1	2.55	0.41
2:D:127:ARG:HG3	2:D:159:HIS:CG	2.55	0.41
1:C:11:ASP:HB2	2:D:140:ILE:HB	2.01	0.41
1:C:141:ARG:O	1:C:143:SER:OG	2.36	0.41
1:C:179:VAL:O	1:C:254:PRO:HB3	2.20	0.41
2:D:74:GLU:HB3	2:D:77:ILE:HG22	2.02	0.41
1:C:130:ILE:HG12	1:C:131:ARG:N	2.35	0.41
1:C:156:SER:HB3	1:C:159:ALA:HB3	2.01	0.41
1:A:237:LEU:HD22	1:A:241:ASP:HB3	2.02	0.41
2:D:67:ASP:HB2	2:D:68:ASN:H	1.73	0.41
1:A:237:LEU:HD23	1:A:237:LEU:HA	1.92	0.41
1:A:302:VAL:HG11	2:B:65:LEU:HD13	2.02	0.40
1:C:67:LEU:HB2	1:C:95:ASP:HB3	2.02	0.40
1:C:11:ASP:O	2:D:140:ILE:N	2.52	0.40
1:C:130:ILE:HD11	1:C:154:LEU:HB3	2.03	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	316/321 (98%)	296 (94%)	16 (5%)	4 (1%)	12	30
1	C	316/321 (98%)	290 (92%)	23 (7%)	3 (1%)	17	40
2	B	167/180 (93%)	156 (93%)	8 (5%)	3 (2%)	8	21
2	D	167/180 (93%)	155 (93%)	8 (5%)	4 (2%)	6	15
All	All	966/1002 (96%)	897 (93%)	55 (6%)	14 (1%)	11	28

All (14) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	12	ASN
1	A	156	SER
2	B	57	GLU
1	C	157	ASN
2	D	57	GLU
1	A	157	ASN
1	A	158(A)	ASP
1	C	143	SER
2	D	12	ASN
1	A	143	SER
2	B	10	ILE
2	D	11	GLU
2	D	10	ILE
1	C	262	GLY

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	267/269 (99%)	253 (95%)	14 (5%)	23	49
1	C	267/269 (99%)	260 (97%)	7 (3%)	46	75

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	145/155 (94%)	137 (94%)	8 (6%)	21	46
2	D	145/155 (94%)	139 (96%)	6 (4%)	30	59
All	All	824/848 (97%)	789 (96%)	35 (4%)	30	58

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	24	THR
1	A	29	LEU
1	A	75	GLN
1	A	78	GLN
1	A	96	VAL
1	A	130	ILE
1	A	156	SER
1	A	158(A)	ASP
1	A	199	ASN
1	A	203	THR
1	A	211	GLN
1	A	237	LEU
1	A	277	CYS
2	B	11	GLU
2	B	52	LEU
2	B	54	ARG
2	B	56	ILE
2	B	64	GLU
2	B	65	LEU
2	B	124	ARG
2	B	144	CYS
1	C	18	HIS
1	C	75	GLN
1	C	78	GLN
1	C	158(A)	ASP
1	C	203	THR
1	C	211	GLN
1	C	237	LEU
2	D	12	ASN
2	D	54	ARG
2	D	64	GLU
2	D	65	LEU
2	D	124	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	129	ASN

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

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](#)

5 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	E	1	1,3	14,14,15	0.66	1 (7%)	17,19,21	0.69	0
3	NAG	E	2	3	14,14,15	0.25	0	17,19,21	0.66	0
3	BMA	E	3	3	11,11,12	0.82	0	15,15,17	0.90	1 (6%)
4	GAL	F	1	4	11,11,12	0.85	0	15,15,17	1.04	1 (6%)
4	SIA	F	2	4	17,20,21	0.42	0	21,28,31	1.32	3 (14%)

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	E	1	1,3	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	E	2	3	-	2/6/23/26	0/1/1/1
3	BMA	E	3	3	-	0/2/19/22	0/1/1/1
4	GAL	F	1	4	-	2/2/19/22	0/1/1/1
4	SIA	F	2	4	-	5/14/34/38	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	1	NAG	O5-C1	-2.28	1.40	1.43

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	2	SIA	C6-O6-C2	4.10	120.11	111.34
4	F	2	SIA	C3-C4-C5	-2.71	108.19	111.46
4	F	2	SIA	C3-C2-C1	-2.40	106.69	111.93
4	F	1	GAL	O3-C3-C4	-2.15	105.37	110.35
3	E	3	BMA	O2-C2-C3	-2.07	105.99	110.14

There are no chirality outliers.

All (9) torsion outliers are listed below:

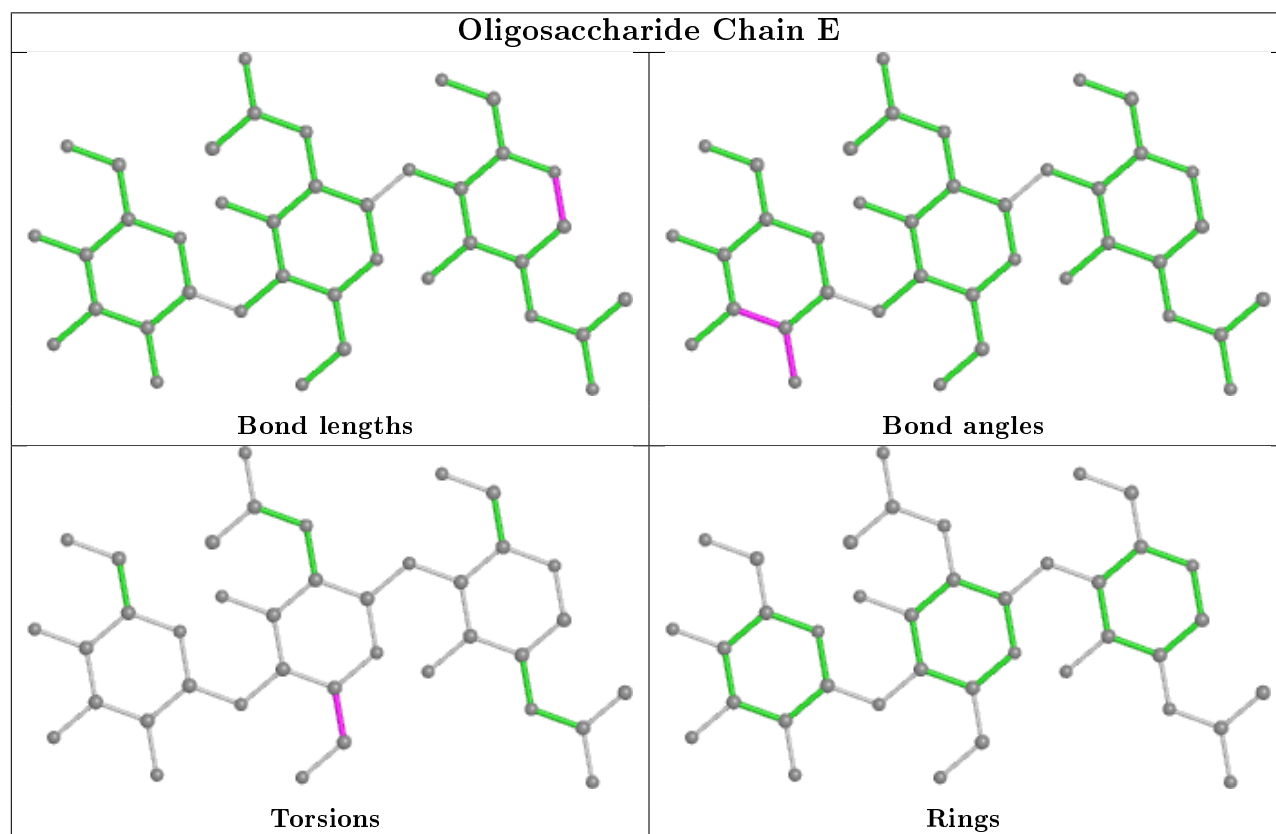
Mol	Chain	Res	Type	Atoms
4	F	2	SIA	C6-C7-C8-O8
4	F	2	SIA	O7-C7-C8-O8
4	F	1	GAL	O5-C5-C6-O6
4	F	1	GAL	C4-C5-C6-O6
4	F	2	SIA	O7-C7-C8-C9
4	F	2	SIA	C6-C7-C8-C9
3	E	2	NAG	O5-C5-C6-O6
3	E	2	NAG	C4-C5-C6-O6
4	F	2	SIA	C6-C5-N5-C10

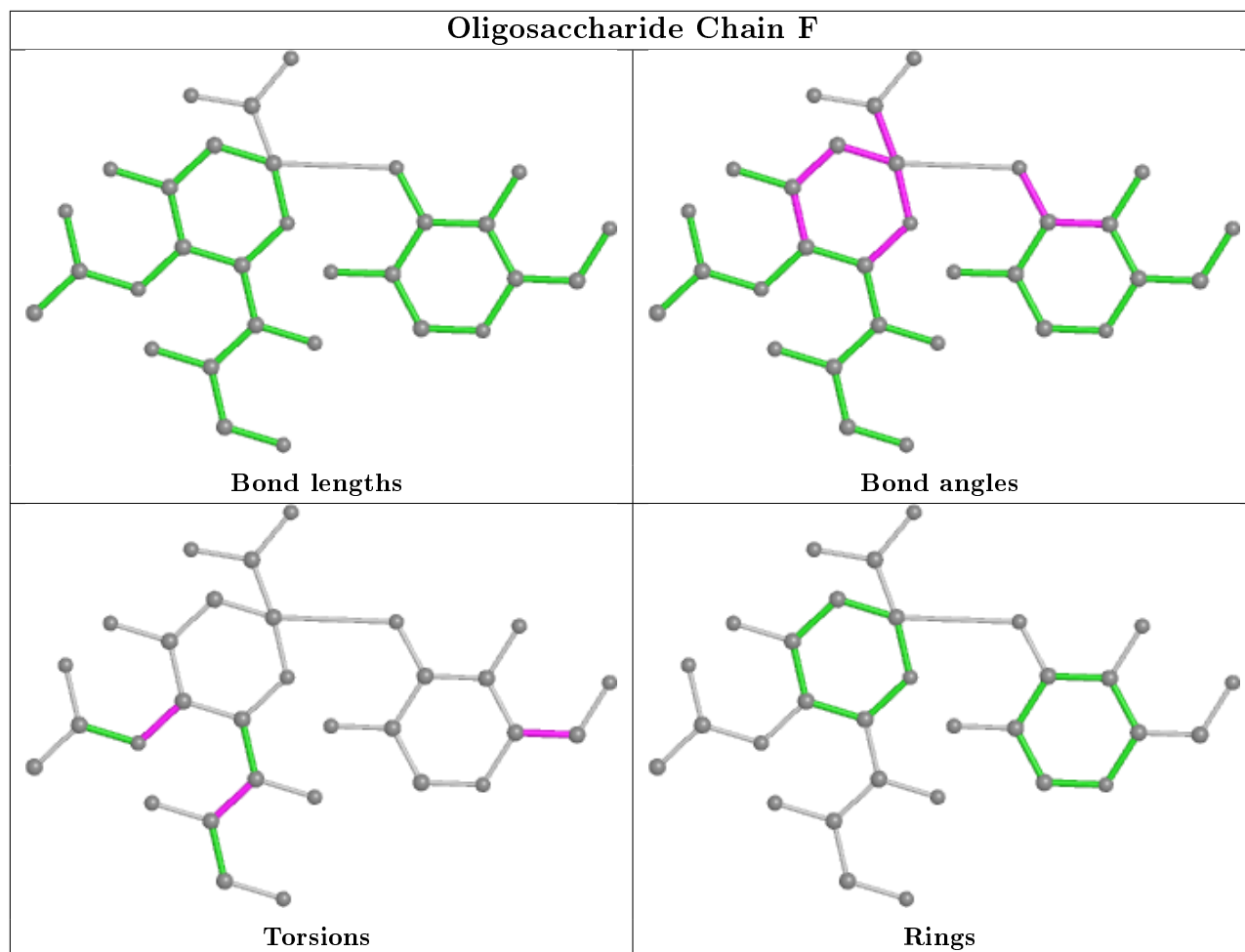
There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	1	GAL	1	0
4	F	2	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.





5.6 Ligand geometry [i](#)

4 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	SIA	C	401	-	17,20,21	0.32	0	21,28,31	0.74	1 (4%)
5	NAG	D	201	2	14,14,15	0.40	0	17,19,21	0.29	0
5	NAG	B	201	2	14,14,15	0.27	0	17,19,21	0.40	0
5	NAG	A	404	1	14,14,15	0.27	0	17,19,21	0.52	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	SIA	C	401	-	-	6/14/34/38	0/1/1/1
5	NAG	D	201	2	-	2/6/23/26	0/1/1/1
5	NAG	B	201	2	-	2/6/23/26	0/1/1/1
5	NAG	A	404	1	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	C	401	SIA	C6-O6-C2	2.39	116.45	111.34

There are no chirality outliers.

All (13) torsion outliers are listed below:

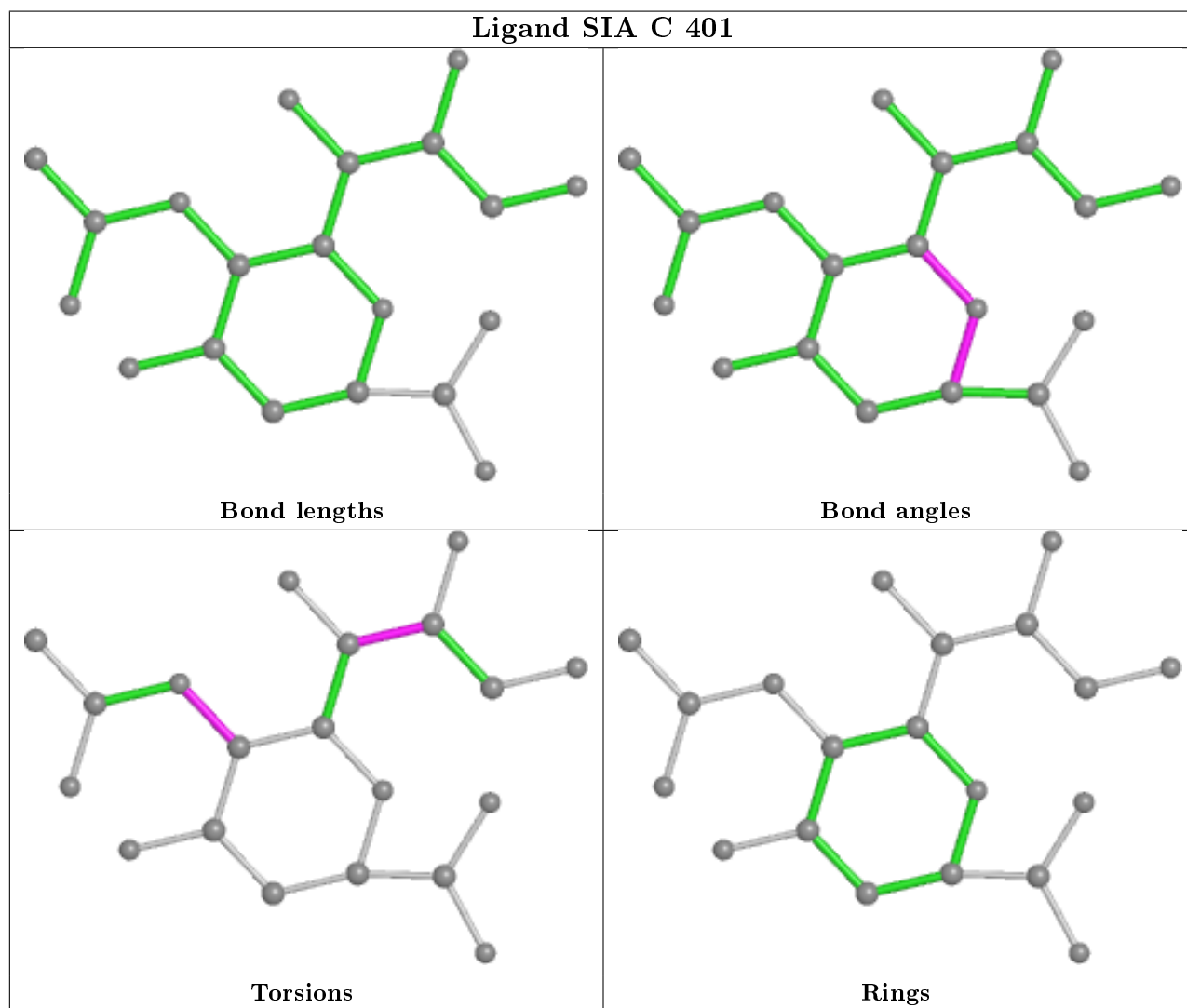
Mol	Chain	Res	Type	Atoms
6	C	401	SIA	C6-C7-C8-C9
6	C	401	SIA	C6-C7-C8-O8
6	C	401	SIA	O7-C7-C8-O8
5	B	201	NAG	C4-C5-C6-O6
5	D	201	NAG	O5-C5-C6-O6
5	D	201	NAG	C4-C5-C6-O6
5	B	201	NAG	O5-C5-C6-O6
6	C	401	SIA	O7-C7-C8-C9
5	A	404	NAG	O5-C5-C6-O6
5	A	404	NAG	C4-C5-C6-O6
6	C	401	SIA	C4-C5-N5-C10
6	C	401	SIA	C6-C5-N5-C10
5	A	404	NAG	C3-C2-N2-C7

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	401	SIA	1	0
5	D	201	NAG	1	0
5	B	201	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 all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



5.7 Other polymers [\(i\)](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data [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	318/321 (99%)	0.03	4 (1%) 77 78	29, 56, 97, 148	0
1	C	318/321 (99%)	0.98	51 (16%) 1 1	63, 115, 192, 283	0
2	B	169/180 (93%)	0.08	1 (0%) 89 91	24, 52, 88, 149	0
2	D	169/180 (93%)	2.02	63 (37%) 0 0	75, 182, 229, 271	0
All	All	974/1002 (97%)	0.69	119 (12%) 4 3	24, 85, 206, 283	0

All (119) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	319	GLY	10.5
1	C	289	ILE	9.1
2	D	7	ALA	9.1
1	C	58	THR	8.8
2	D	138	PHE	8.7
2	D	23	GLY	8.5
2	D	36	ALA	8.4
2	D	35	ALA	7.9
2	D	24	PHE	7.4
2	D	126	LEU	6.9
2	D	130	ALA	6.9
1	C	13	ILE	6.7
2	D	122	VAL	6.5
2	D	52	LEU	6.4
2	D	140	ILE	6.2
1	C	15	LEU	6.0
1	A	327	PRO	5.8
1	C	37	VAL	5.5
2	D	32	GLU	5.5
2	D	51	LYS	5.3
1	C	86	LEU	5.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	143	LYS	5.0
2	D	119	TYR	5.0
1	C	327	PRO	5.0
1	C	33	GLY	4.9
2	D	131	GLU	4.8
2	D	33	GLY	4.8
2	D	6	ILE	4.7
2	D	157	TYR	4.6
2	D	172	GLN	4.4
2	D	171	ILE	4.3
2	D	48	ILE	4.3
2	D	158	ASP	4.2
2	D	26	HIS	4.2
1	C	277	CYS	4.2
1	C	224	ASN	4.1
2	D	16	GLY	4.0
2	D	55	LEU	4.0
1	C	12	LYS	4.0
2	D	100	VAL	4.0
2	D	25	ARG	3.9
1	C	20	VAL	3.9
1	C	25	LYS	3.9
1	C	288	ILE	3.9
2	D	34	THR	3.9
1	C	104	ASN	3.8
1	C	140	ARG	3.7
2	D	168	GLN	3.7
2	D	99	LEU	3.6
2	D	170	ARG	3.6
1	C	226	LEU	3.6
2	D	8	GLY	3.5
1	C	59	VAL	3.5
1	C	297	ILE	3.4
1	C	22	ASN	3.4
2	D	141	PHE	3.4
2	D	103	GLU	3.4
1	C	274	VAL	3.3
1	C	21	SER	3.3
1	C	279	GLY	3.2
2	D	144	CYS	3.2
2	D	127	ARG	3.2
2	D	18	ILE	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	11	ASP	3.1
1	C	318	THR	3.1
2	D	118	LEU	3.0
2	D	44	ALA	3.0
1	C	326	ILE	3.0
2	D	121	ARG	2.9
2	D	139	GLU	2.9
2	D	59	THR	2.9
1	C	276(A)	ASN	2.8
2	D	142	HIS	2.8
1	C	320	MET	2.8
1	C	323	VAL	2.7
2	D	124	ARG	2.7
2	D	60	ASN	2.7
1	C	46	THR	2.7
1	C	294	PHE	2.6
1	C	57	ARG	2.5
1	C	82	PHE	2.5
2	D	113	SER	2.5
2	D	42	GLN	2.5
1	C	276	ALA	2.5
2	D	125	GLN	2.5
1	C	231	ASP	2.5
2	B	11	GLU	2.4
1	C	199	ASN	2.4
2	D	45	ILE	2.4
2	D	152	ILE	2.4
2	D	22	TYR	2.4
1	C	84	ALA	2.4
2	D	132	GLU	2.4
1	A	231	ASP	2.4
2	D	53	ASN	2.4
2	D	129	ASN	2.4
1	C	88	ILE	2.4
1	C	54	LYS	2.3
1	C	16	GLY	2.3
1	C	45	ARG	2.3
2	D	74	GLU	2.3
2	D	107	THR	2.3
1	C	275	ASP	2.3
1	C	51	ILE	2.3
1	A	199	ASN	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	31	GLY	2.2
2	D	117	LYS	2.2
2	D	58	LYS	2.2
1	C	112	LEU	2.2
1	C	40	THR	2.2
1	C	278	GLU	2.1
2	D	102	MET	2.1
1	C	24	THR	2.1
1	C	207	SER	2.1
2	D	5	ALA	2.1
2	D	92	TRP	2.0
1	C	314	LEU	2.0
1	C	38	ASN	2.0
1	A	212	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](#)

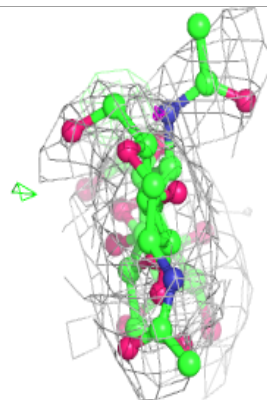
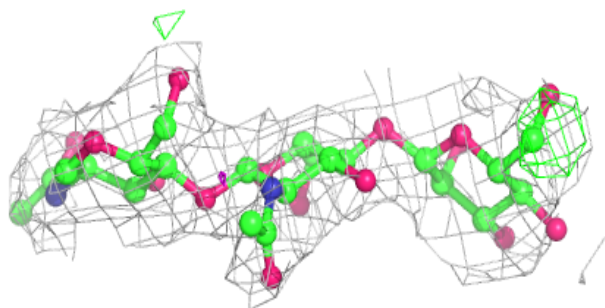
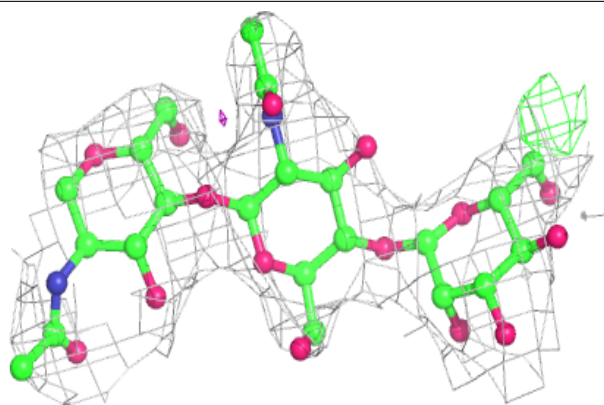
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	GAL	F	1	11/12	0.70	0.37	167,182,195,196	0
3	BMA	E	3	11/12	0.74	0.16	151,153,154,156	0
3	NAG	E	2	14/15	0.80	0.25	120,126,134,141	0
4	SIA	F	2	20/21	0.84	0.28	81,98,109,111	0
3	NAG	E	1	14/15	0.91	0.15	74,90,97,110	0

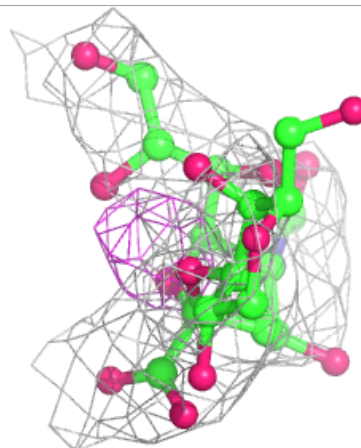
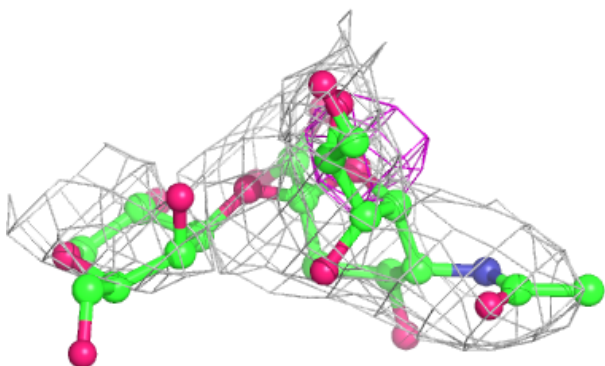
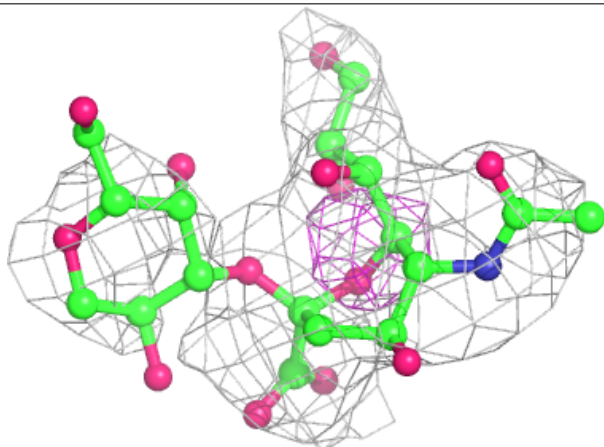
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain E:

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

**Electron density around Chain F:**

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

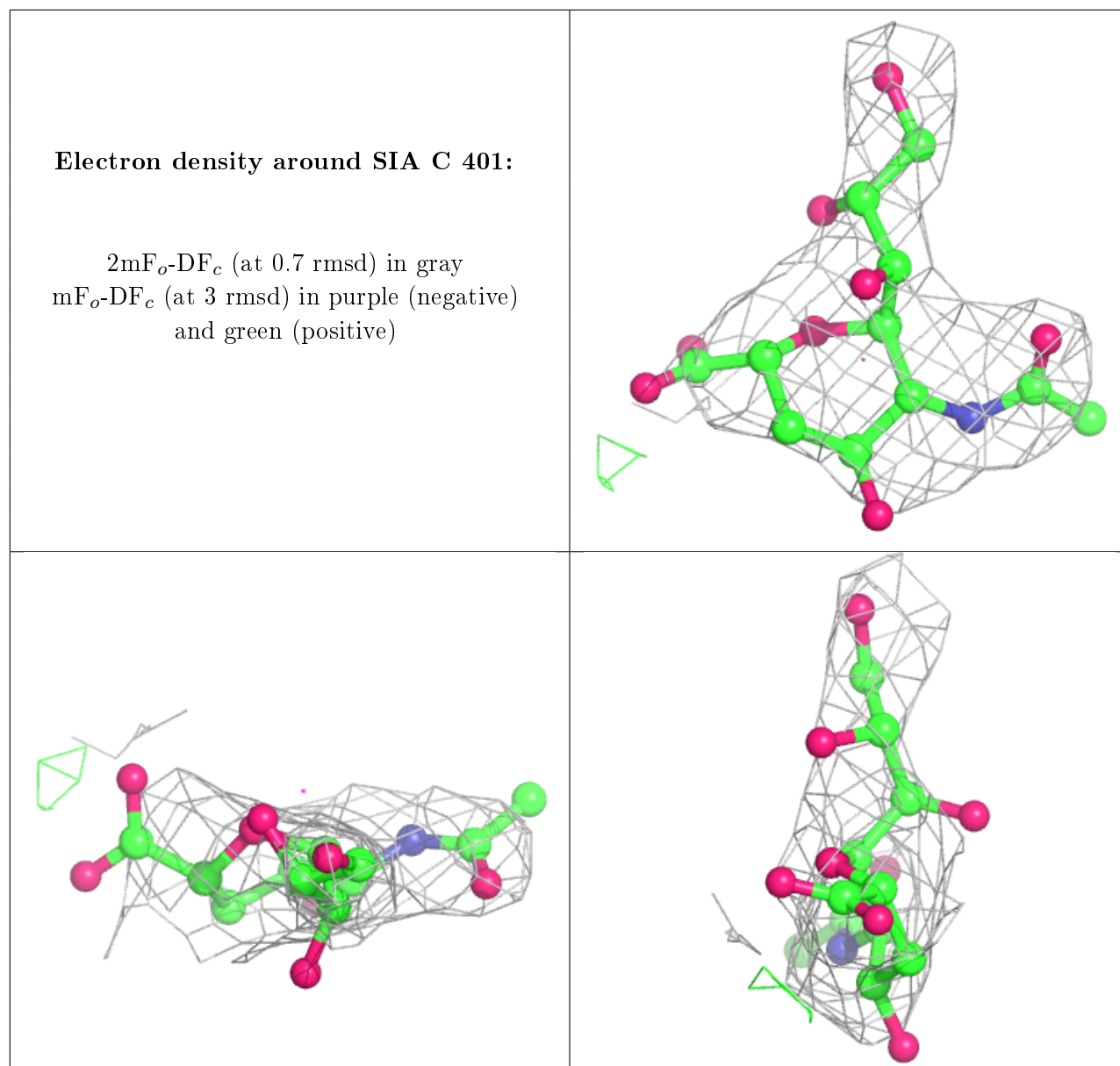


6.4 Ligands

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
5	NAG	D	201	14/15	0.77	0.31	142,143,144,144	0
6	SIA	C	401	20/21	0.78	0.40	113,122,132,132	0
5	NAG	A	404	14/15	0.79	0.20	123,134,145,147	0
5	NAG	B	201	14/15	0.90	0.19	61,74,79,80	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



6.5 Other polymers [i](#)

There are no such residues in this entry.