



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 11, 2024 – 08:31 AM JST

PDB ID : 4YY7  
Title : The structure of hemagglutinin from a H6N1 influenza virus (A/chicken/Taiwan/A2837/2013) in complex with avian receptor analog 3'SLNLN  
Authors : Wang, F.; Qi, J.; Bi, Y.; Zhang, W.; Wang, M.; Wang, M.; Liu, J.; Yan, J.; Shi, Y.; Gao, G.F.  
Deposited on : 2015-03-23  
Resolution : 2.99 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

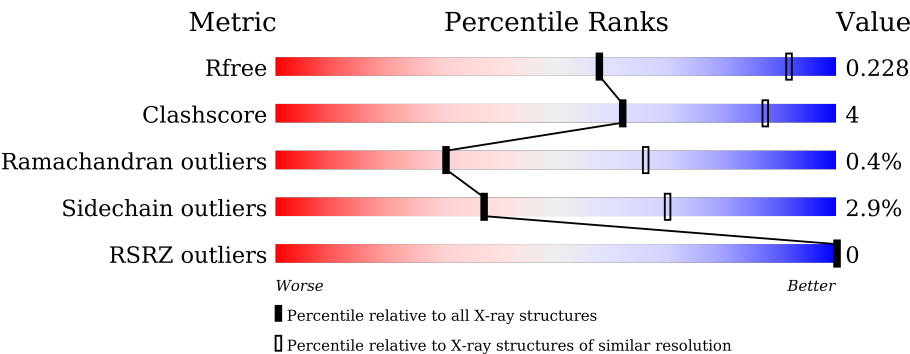
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

# 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 2.99 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))
$R_{free}$	164625	2511 (3.00-3.00)
Clashscore	180529	2866 (3.00-3.00)
Ramachandran outliers	177936	2778 (3.00-3.00)
Sidechain outliers	177891	2781 (3.00-3.00)
RSRZ outliers	164620	2523 (3.00-3.00)

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  $\geq 3$ , 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	325	<div><div></div><div>86%13%.</div></div>
1	C	325	<div><div></div><div>86%13%.</div></div>
2	B	164	<div><div></div><div>89%11%</div></div>
2	D	164	<div><div></div><div>94%6%</div></div>
3	E	2	<div><div></div><div>50%50%</div></div>
3	G	2	<div><div></div><div>50%50%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain	
4	F	3	<div><div></div></div> 33%	<div><div></div></div> 67%
4	H	3	<div><div></div></div> 33%	<div><div></div></div> 67%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7972 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 HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	325	Total	C	N	O	S	0	0	0
			2563	1625	436	489	13			
1	C	325	Total	C	N	O	S	0	0	0
			2563	1625	436	489	13			

- Molecule 2 is a protein called HA2.

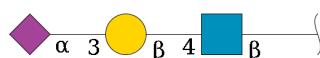
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	164	Total	C	N	O	S	0	0	0
			1321	825	229	260	7			
2	D	164	Total	C	N	O	S	0	0	0
			1321	825	229	260	7			

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



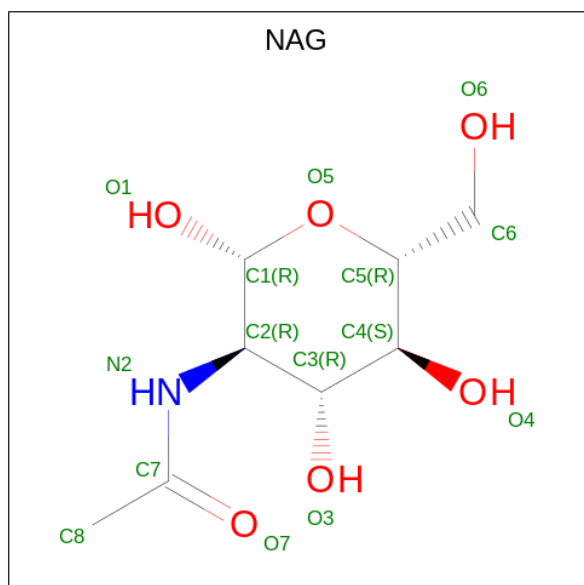
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	G	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



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

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

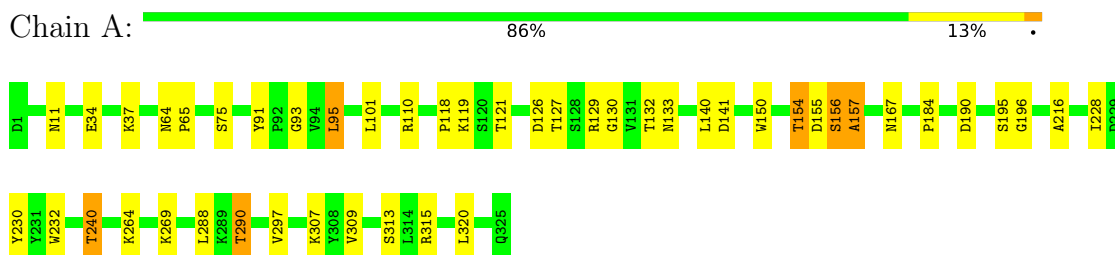


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	C	1	Total	C	N	O	0	0
			14	8	1	5		
5	D	1	Total	C	N	O	0	0
			14	8	1	5		

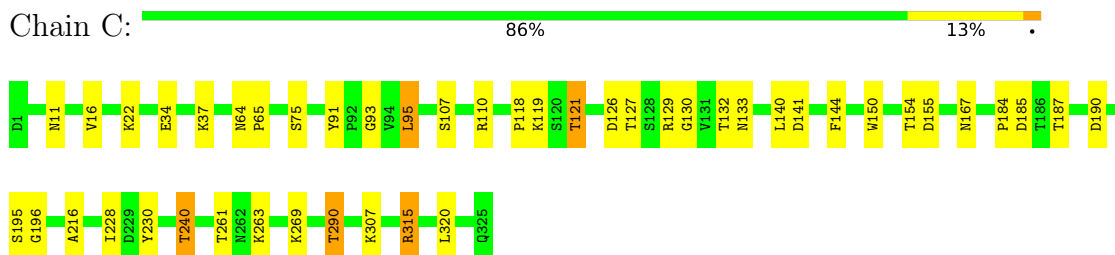
### 3 Residue-property plots

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.

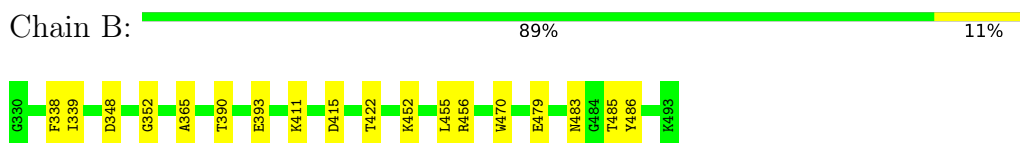
- Molecule 1: HA1



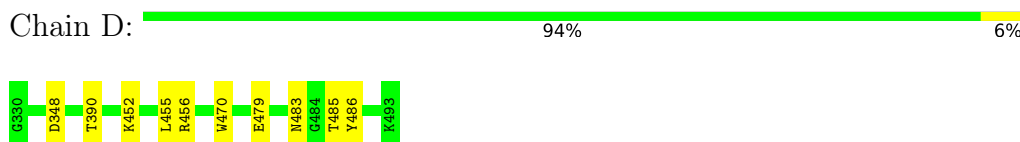
- Molecule 1: HA1



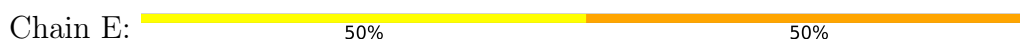
- Molecule 2: HA2



- Molecule 2: HA2



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





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



- Molecule 4: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	96.77Å 96.77Å 132.51Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.45 – 2.99 45.45 – 2.99	Depositor EDS
% Data completeness (in resolution range)	98.7 (45.45-2.99) 98.7 (45.45-2.99)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.17 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.8.4_1496	Depositor
R, $R_{free}$	0.199 , 0.226 0.203 , 0.228	Depositor DCC
$R_{free}$ test set	1390 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	79.0	Xtriage
Anisotropy	0.366	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 39.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.020 for -h,-k,l 0.449 for h,-h-k,-l 0.023 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	7972	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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, NAG, GAL

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	0/2625	0.46	0/3571
1	C	0.26	0/2625	0.46	0/3571
2	B	0.29	0/1349	0.46	0/1818
2	D	0.29	0/1349	0.46	0/1818
All	All	0.27	0/7948	0.46	0/10778

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	2563	0	2508	23	0
1	C	2563	0	2508	23	0
2	B	1321	0	1229	13	0
2	D	1321	0	1229	8	0
3	E	28	0	25	1	0
3	G	28	0	25	2	0
4	F	46	0	40	0	0
4	H	46	0	40	0	0
5	A	14	0	13	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	14	0	13	3	0
5	C	14	0	13	1	0
5	D	14	0	13	3	0
All	All	7972	0	7656	63	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 4.

All (63) 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:D:452:LYS:O	2:D:455:LEU:O	1.90	0.89
1:C:167:ASN:HA	1:C:240:THR:HB	1.59	0.83
2:D:485:THR:HG21	5:D:501:NAG:H82	1.68	0.76
1:A:167:ASN:HA	1:A:240:THR:HB	1.66	0.75
2:B:452:LYS:O	2:B:455:LEU:O	2.11	0.68
2:B:455:LEU:HD23	2:B:486:TYR:CE2	2.29	0.68
1:A:126:ASP:OD2	1:A:129:ARG:NH1	2.27	0.67
1:C:240:THR:HG21	3:G:1:NAG:H82	1.75	0.67
1:A:154:THR:O	1:A:156:SER:N	2.30	0.63
1:C:22:LYS:HD2	5:C:601:NAG:H82	1.82	0.60
1:C:130:GLY:HA3	1:C:150:TRP:HB3	1.83	0.59
2:B:485:THR:HG21	5:B:501:NAG:H82	1.83	0.59
1:C:16:VAL:HG12	1:C:315:ARG:HG2	1.88	0.55
2:D:479:GLU:HG2	5:D:501:NAG:O6	2.08	0.53
1:C:240:THR:CG2	3:G:1:NAG:H82	2.39	0.52
2:D:455:LEU:HD23	2:D:486:TYR:CE2	2.45	0.51
1:C:126:ASP:OD2	1:C:129:ARG:NH2	2.43	0.49
1:A:195:SER:OG	1:A:196:GLY:N	2.44	0.49
1:A:240:THR:HG21	3:E:1:NAG:H82	1.95	0.48
2:D:348:ASP:N	2:D:348:ASP:OD1	2.46	0.48
2:B:348:ASP:N	2:B:348:ASP:OD1	2.47	0.47
2:D:483:ASN:ND2	5:D:501:NAG:O7	2.48	0.47
1:C:11:ASN:OD1	1:C:11:ASN:N	2.48	0.46
1:A:307:LYS:HE3	2:B:390:THR:O	2.14	0.46
2:B:483:ASN:ND2	5:B:501:NAG:O7	2.49	0.46
1:A:93:GLY:HA3	1:A:228:ILE:O	2.16	0.45
1:A:91:TYR:HD2	1:A:132:THR:HG21	1.82	0.45
1:A:101:LEU:HB2	1:A:232:TRP:CE2	2.52	0.44
1:A:288:LEU:HD21	1:A:297:VAL:HG21	1.98	0.44
2:B:479:GLU:HG2	5:B:501:NAG:O6	2.17	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:156:SER:OG	1:A:157:ALA:N	2.51	0.44
1:A:130:GLY:HA3	1:A:150:TRP:HB3	1.99	0.43
1:C:34:GLU:HB2	1:C:290:THR:HG21	1.99	0.43
1:C:307:LYS:HE3	2:D:390:THR:O	2.18	0.43
1:C:95:LEU:HA	1:C:230:TYR:HB2	2.00	0.43
1:C:119:LYS:HD3	1:C:127:THR:HB	2.01	0.43
2:B:411:LYS:HG2	2:B:415:ASP:OD2	2.19	0.43
1:A:34:GLU:HB2	1:A:290:THR:HG21	2.00	0.42
1:A:184:PRO:HA	1:A:216:ALA:O	2.19	0.42
1:C:65:PRO:HG3	1:C:144:PHE:O	2.19	0.42
1:A:95:LEU:HA	1:A:230:TYR:HB2	2.00	0.42
1:A:118:PRO:O	1:A:121:THR:OG1	2.28	0.42
1:C:195:SER:OG	1:C:196:GLY:N	2.52	0.42
1:A:264:LYS:NZ	2:B:393:GLU:OE1	2.45	0.42
1:A:75:SER:OG	1:A:110:ARG:HD2	2.19	0.42
2:D:470:TRP:HA	2:D:470:TRP:CE3	2.54	0.41
1:C:93:GLY:HA3	1:C:228:ILE:O	2.20	0.41
1:C:185:ASP:OD1	1:C:187:THR:OG1	2.34	0.41
2:B:338:PHE:CE1	2:B:339:ILE:HG13	2.55	0.41
1:C:118:PRO:O	1:C:121:THR:OG1	2.29	0.41
1:A:11:ASN:OD1	1:A:11:ASN:N	2.43	0.41
1:C:64:ASN:HA	1:C:65:PRO:HD3	1.90	0.41
1:C:133:ASN:HA	1:C:141:ASP:O	2.21	0.41
1:A:119:LYS:HD3	1:A:127:THR:HB	2.03	0.41
1:C:184:PRO:HA	1:C:216:ALA:O	2.21	0.41
1:A:133:ASN:HA	1:A:141:ASP:O	2.21	0.41
2:B:352:GLY:HA3	2:B:365:ALA:HA	2.03	0.41
2:B:470:TRP:HA	2:B:470:TRP:CE3	2.55	0.41
1:A:309:VAL:CG2	2:B:422:THR:HA	2.51	0.40
1:C:91:TYR:HD2	1:C:132:THR:HG21	1.86	0.40
1:C:107:SER:O	1:C:261:THR:HG22	2.21	0.40
1:A:64:ASN:HA	1:A:65:PRO:HD3	1.90	0.40
1:C:75:SER:OG	1:C:110:ARG:HD2	2.20	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	323/325 (99%)	300 (93%)	20 (6%)	3 (1%)	14	49
1	C	323/325 (99%)	305 (94%)	17 (5%)	1 (0%)	37	70
2	B	162/164 (99%)	144 (89%)	18 (11%)	0	100	100
2	D	162/164 (99%)	146 (90%)	16 (10%)	0	100	100
All	All	970/978 (99%)	895 (92%)	71 (7%)	4 (0%)	30	66

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	155	ASP
1	A	156	SER
1	C	155	ASP
1	A	157	ALA

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	285/285 (100%)	274 (96%)	11 (4%)	27	61
1	C	285/285 (100%)	273 (96%)	12 (4%)	25	59
2	B	139/139 (100%)	138 (99%)	1 (1%)	81	91
2	D	139/139 (100%)	138 (99%)	1 (1%)	81	91
All	All	848/848 (100%)	823 (97%)	25 (3%)	37	70

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	LYS
1	A	95	LEU
1	A	140	LEU
1	A	154	THR
1	A	190	ASP
1	A	240	THR
1	A	269	LYS
1	A	290	THR
1	A	313	SER
1	A	315	ARG
1	A	320	LEU
2	B	456	ARG
1	C	37	LYS
1	C	95	LEU
1	C	121	THR
1	C	140	LEU
1	C	154	THR
1	C	190	ASP
1	C	240	THR
1	C	263	LYS
1	C	269	LYS
1	C	290	THR
1	C	315	ARG
1	C	320	LEU
2	D	456	ARG

Sometimes 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 ⓘ

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

## 5.5 Carbohydrates ⓘ

10 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	3,1	14,14,15	1.07	1 (7%)	17,19,21	1.56	1 (5%)
3	NAG	E	2	3	14,14,15	1.37	1 (7%)	17,19,21	0.96	1 (5%)
4	NAG	F	1	4	15,15,15	0.13	0	21,21,21	0.18	0
4	GAL	F	2	4	11,11,12	1.32	3 (27%)	15,15,17	1.34	2 (13%)
4	SIA	F	3	4	20,20,21	1.62	4 (20%)	24,28,31	1.93	7 (29%)
3	NAG	G	1	3,1	14,14,15	0.39	0	17,19,21	1.01	1 (5%)
3	NAG	G	2	3	14,14,15	0.62	0	17,19,21	0.58	0
4	NAG	H	1	4	15,15,15	0.15	0	21,21,21	0.21	0
4	GAL	H	2	4	11,11,12	1.36	3 (27%)	15,15,17	1.50	2 (13%)
4	SIA	H	3	4	20,20,21	0.91	1 (5%)	24,28,31	1.19	1 (4%)

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	3,1	-	2/6/23/26	0/1/1/1
3	NAG	E	2	3	-	1/6/23/26	0/1/1/1
4	NAG	F	1	4	-	2/6/26/26	0/1/1/1
4	GAL	F	2	4	-	1/2/19/22	0/1/1/1
4	SIA	F	3	4	-	2/18/34/38	0/1/1/1
3	NAG	G	1	3,1	-	2/6/23/26	0/1/1/1
3	NAG	G	2	3	-	2/6/23/26	0/1/1/1
4	NAG	H	1	4	-	2/6/26/26	0/1/1/1
4	GAL	H	2	4	-	1/2/19/22	0/1/1/1
4	SIA	H	3	4	-	6/18/34/38	0/1/1/1

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	2	NAG	O5-C1	4.60	1.51	1.43
4	F	3	SIA	C2-C1	-4.17	1.48	1.52
3	E	1	NAG	O5-C1	3.66	1.49	1.43
4	F	3	SIA	O6-C2	-3.40	1.39	1.43
4	H	2	GAL	C1-C2	3.04	1.59	1.52
4	F	2	GAL	C1-C2	2.69	1.58	1.52
4	H	3	SIA	O6-C2	-2.67	1.40	1.43
4	F	3	SIA	C5-N5	-2.39	1.42	1.45
4	F	2	GAL	C2-C3	2.35	1.56	1.52
4	H	2	GAL	C2-C3	2.21	1.55	1.52
4	F	2	GAL	O3-C3	2.13	1.48	1.43
4	F	3	SIA	O1B-C1	-2.06	1.23	1.30
4	H	2	GAL	O3-C3	2.02	1.47	1.43

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	1	NAG	C1-O5-C5	5.37	119.47	112.19
4	H	3	SIA	O6-C2-C3	4.11	116.11	110.46
4	F	3	SIA	O6-C2-C3	4.04	116.02	110.46
4	H	2	GAL	C1-C2-C3	3.77	114.30	109.67
4	F	3	SIA	C4-C3-C2	-3.57	103.40	109.81
4	F	3	SIA	O9-C9-C8	-3.52	103.41	111.07
3	G	1	NAG	C1-O5-C5	3.46	116.88	112.19
4	F	2	GAL	C1-C2-C3	3.28	113.70	109.67
4	F	3	SIA	O6-C2-C1	3.28	114.13	107.70
3	E	2	NAG	C1-O5-C5	2.88	116.09	112.19
4	H	2	GAL	C1-O5-C5	2.75	115.92	112.19
4	F	3	SIA	C4-C5-N5	-2.66	105.10	110.38
4	F	3	SIA	C8-C7-C6	-2.45	108.39	113.03
4	F	3	SIA	C6-C5-N5	-2.27	107.15	110.91
4	F	2	GAL	C1-O5-C5	2.15	115.11	112.19

There are no chirality outliers.

All (21) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	F	3	SIA	O8-C8-C9-O9
3	E	1	NAG	C4-C5-C6-O6
3	E	1	NAG	O5-C5-C6-O6
3	G	1	NAG	O5-C5-C6-O6
4	F	3	SIA	C7-C8-C9-O9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	H	3	SIA	C11-C10-N5-C5
4	H	1	NAG	O5-C5-C6-O6
4	F	1	NAG	O5-C5-C6-O6
3	G	1	NAG	C4-C5-C6-O6
4	H	3	SIA	O8-C8-C9-O9
4	H	3	SIA	O10-C10-N5-C5
4	H	1	NAG	C4-C5-C6-O6
4	H	2	GAL	O5-C5-C6-O6
4	F	2	GAL	O5-C5-C6-O6
4	F	1	NAG	C4-C5-C6-O6
3	G	2	NAG	C4-C5-C6-O6
4	H	3	SIA	C7-C8-C9-O9
3	E	2	NAG	O5-C5-C6-O6
3	G	2	NAG	O5-C5-C6-O6
4	H	3	SIA	O1B-C1-C2-O6
4	H	3	SIA	C6-C7-C8-O8

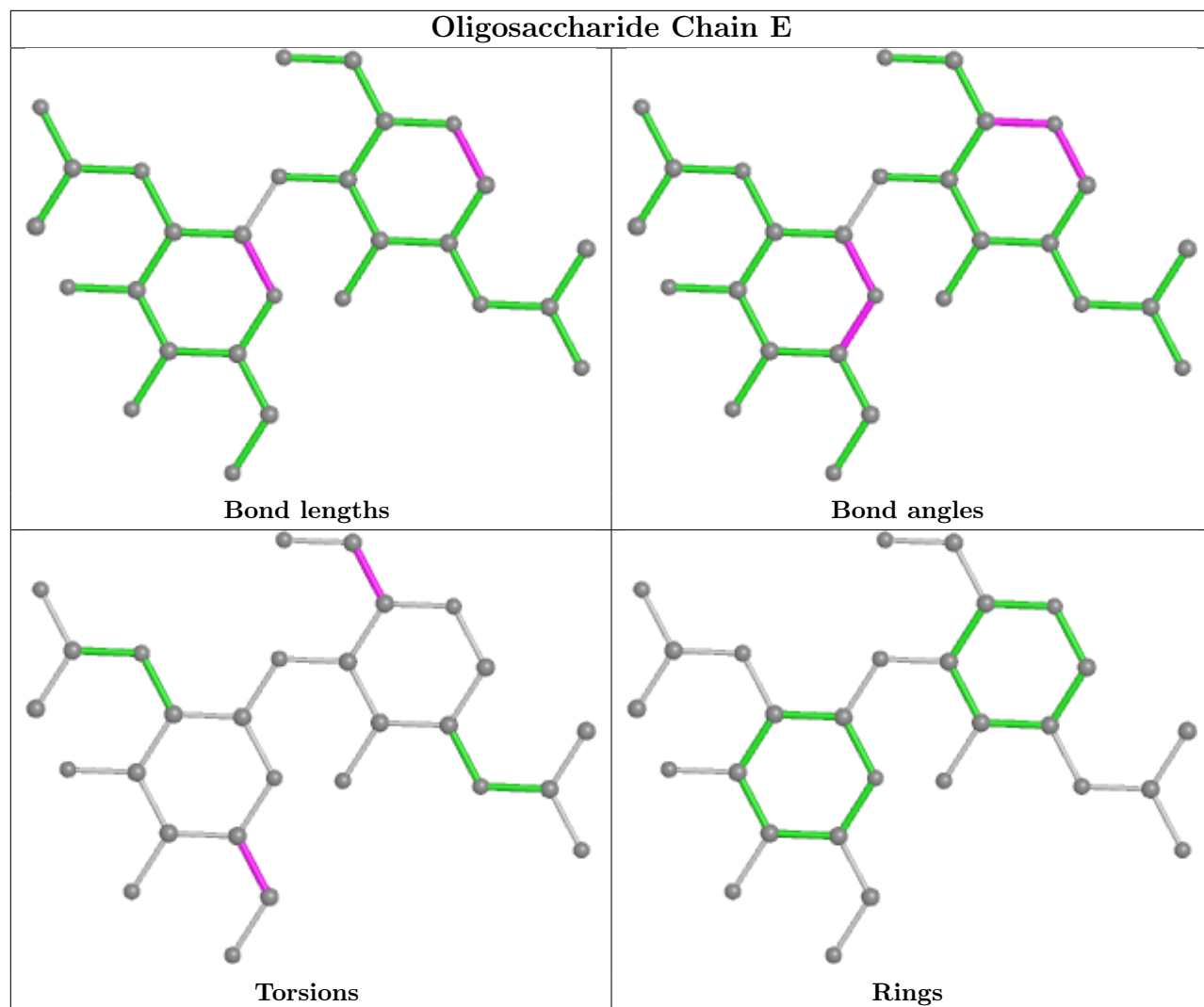
There are no ring outliers.

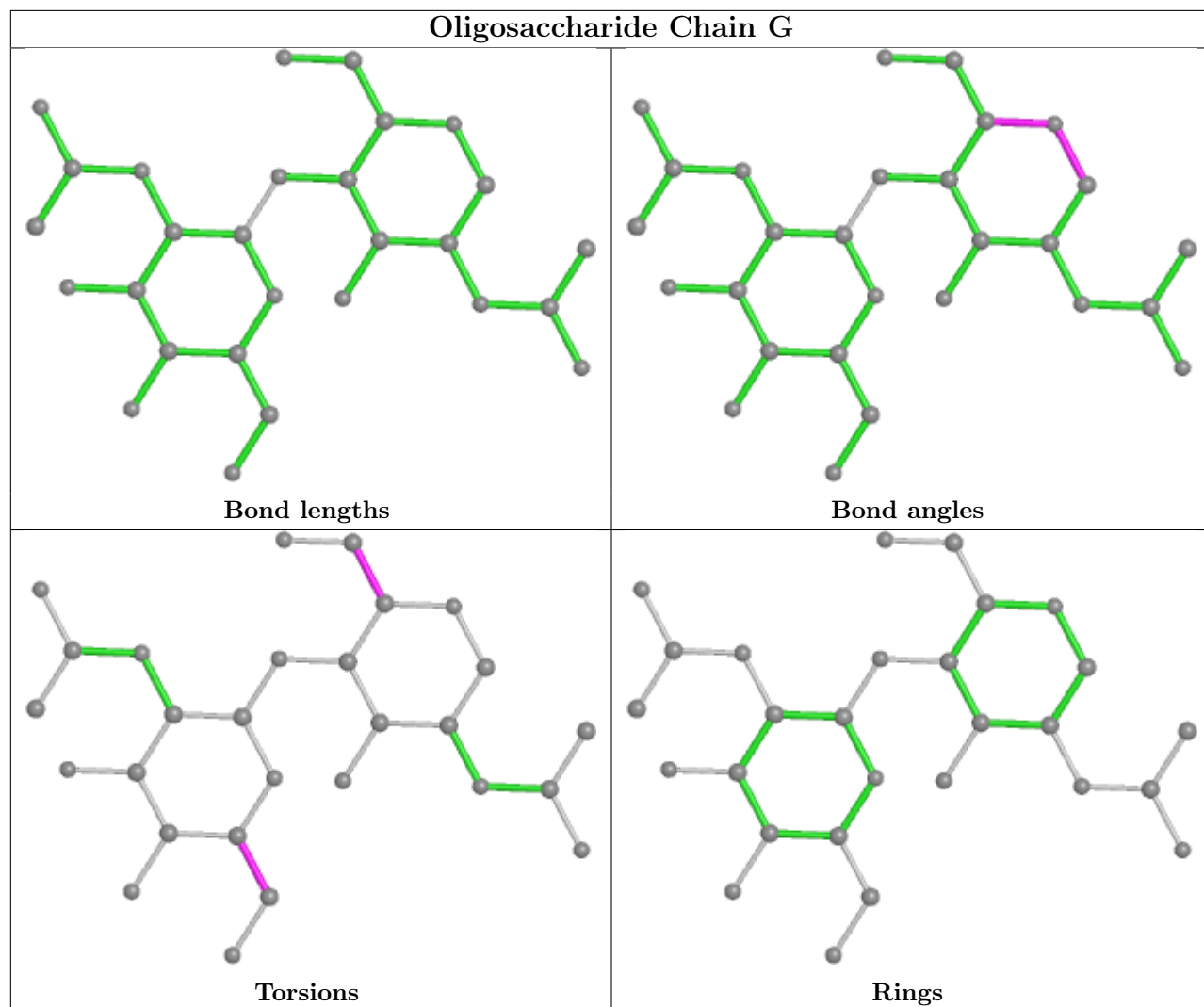
2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	1	NAG	1	0
3	G	1	NAG	2	0

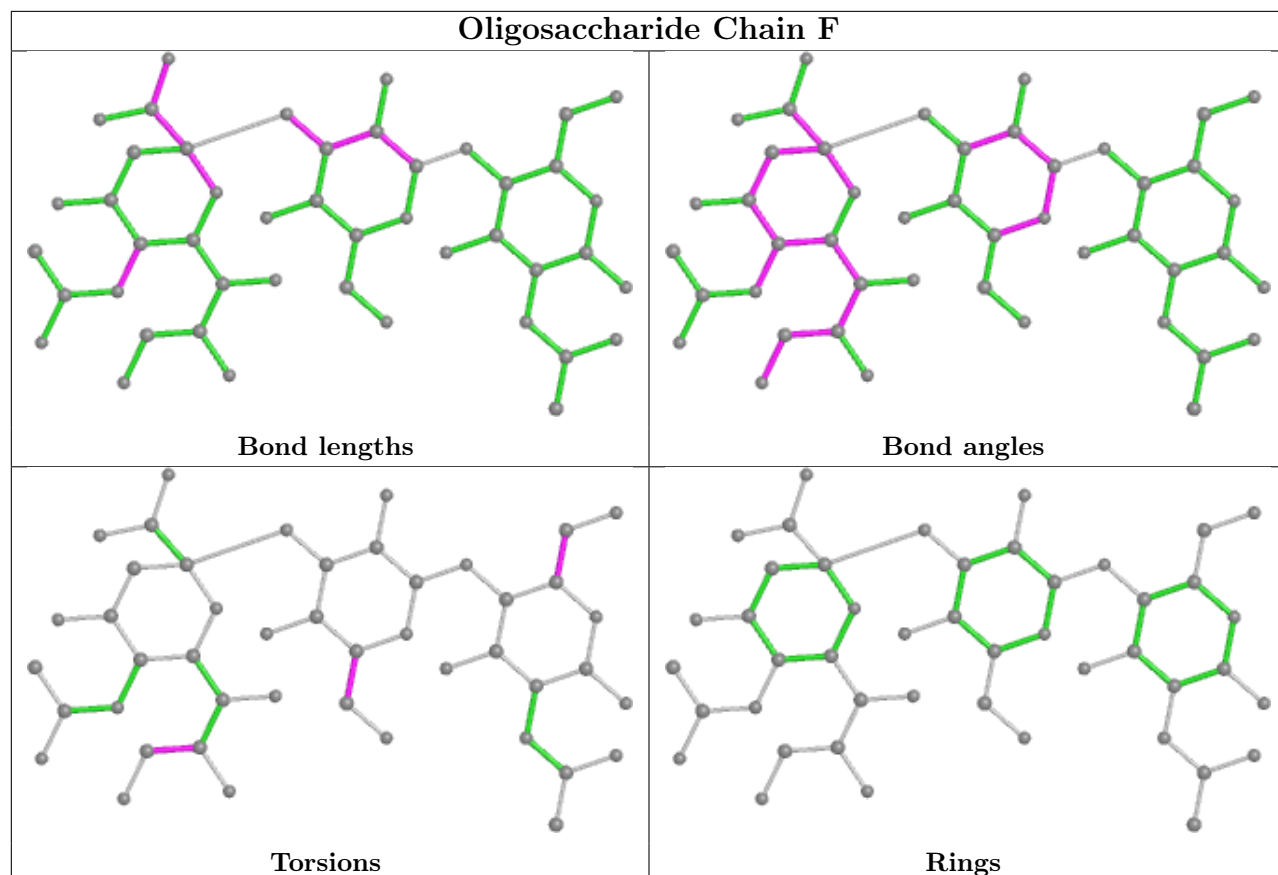
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



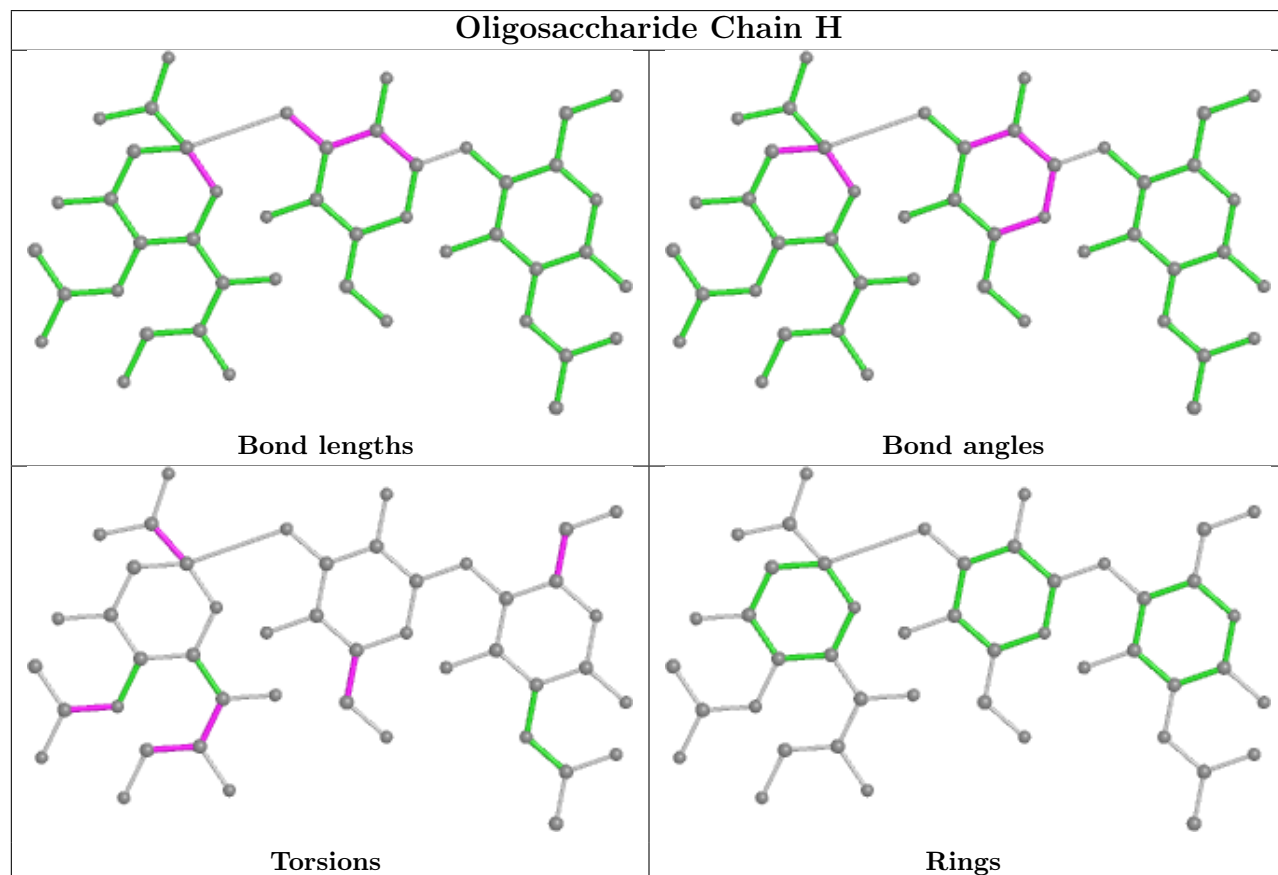




## Oligosaccharide Chain F



## Oligosaccharide Chain H



## 5.6 Ligand geometry

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
5	NAG	D	501	2	14,14,15	1.99	1 (7%)	17,19,21	1.61	1 (5%)
5	NAG	A	401	1	14,14,15	1.29	1 (7%)	17,19,21	0.82	1 (5%)
5	NAG	B	501	2	14,14,15	2.32	2 (14%)	17,19,21	1.37	1 (5%)
5	NAG	C	601	1	14,14,15	1.93	1 (7%)	17,19,21	1.61	1 (5%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	D	501	2	-	0/6/23/26	0/1/1/1
5	NAG	A	401	1	-	0/6/23/26	0/1/1/1
5	NAG	B	501	2	-	0/6/23/26	0/1/1/1
5	NAG	C	601	1	-	2/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	501	NAG	O5-C1	8.18	1.56	1.43
5	D	501	NAG	O5-C1	7.17	1.55	1.43
5	C	601	NAG	O5-C1	6.86	1.54	1.43
5	A	401	NAG	O5-C1	4.33	1.50	1.43
5	B	501	NAG	C1-C2	2.43	1.56	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	601	NAG	C1-O5-C5	5.84	120.11	112.19
5	D	501	NAG	C1-O5-C5	5.74	119.96	112.19

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	501	NAG	C1-O5-C5	4.76	118.64	112.19
5	A	401	NAG	C1-O5-C5	2.93	116.17	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	601	NAG	O5-C5-C6-O6
5	C	601	NAG	C4-C5-C6-O6

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	501	NAG	3	0
5	B	501	NAG	3	0
5	C	601	NAG	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	325/325 (100%)	-1.62	0 100 100	51, 76, 121, 180	0
1	C	325/325 (100%)	-1.60	0 100 100	50, 76, 124, 192	0
2	B	164/164 (100%)	-1.37	0 100 100	50, 115, 171, 207	0
2	D	164/164 (100%)	-1.35	0 100 100	49, 115, 177, 208	0
All	All	978/978 (100%)	-1.52	0 100 100	49, 83, 161, 208	0

There are no RSRZ outliers to report.

### 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<sup>th</sup> 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.

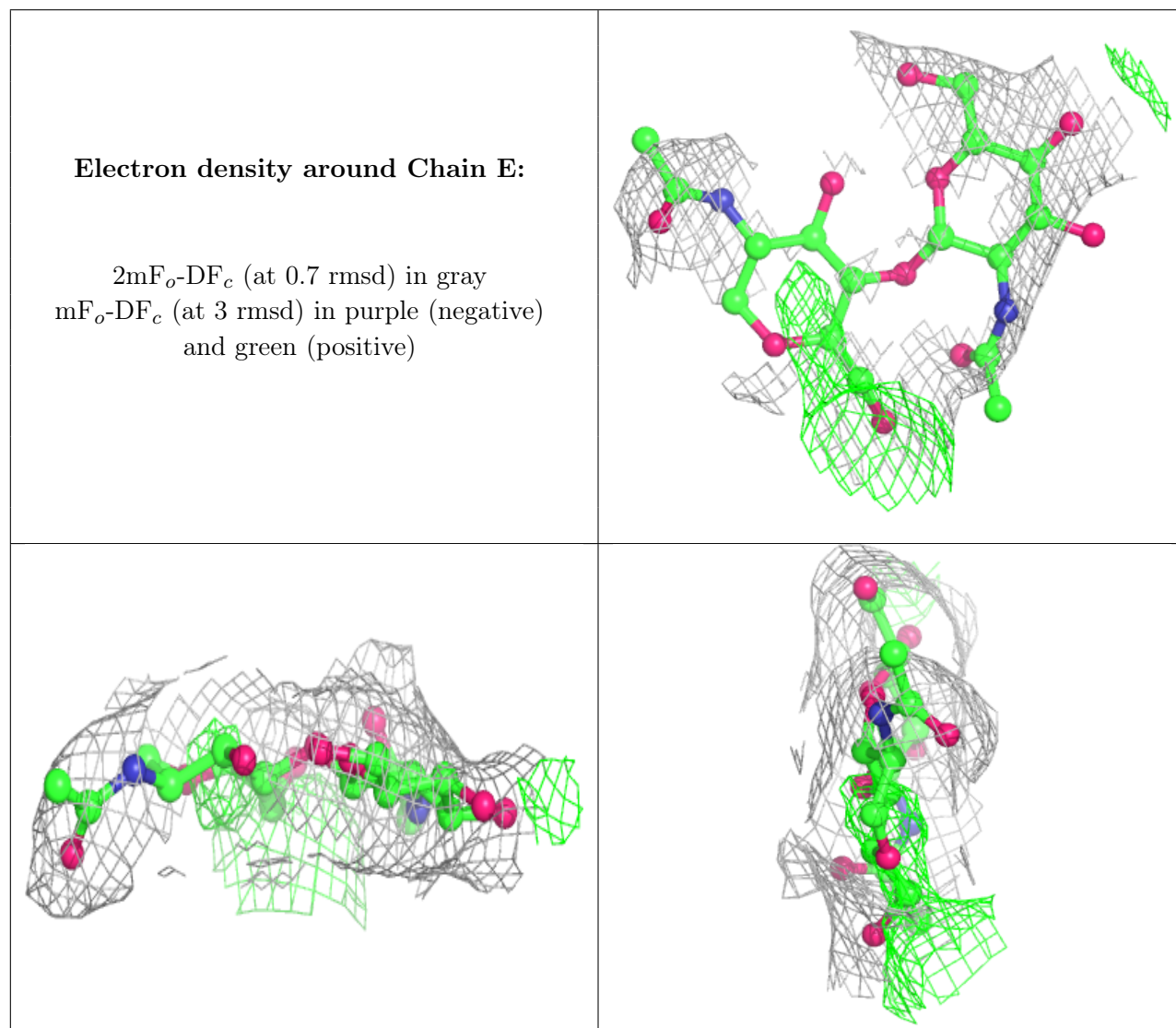
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAG	E	2	14/15	0.97	0.05	120,124,127,131	0
3	NAG	G	2	14/15	0.97	0.06	124,129,132,136	0
3	NAG	G	1	14/15	0.98	0.04	90,96,100,101	0
3	NAG	E	1	14/15	0.98	0.05	88,100,107,107	0
4	NAG	F	1	15/15	0.98	0.06	130,133,136,136	0
4	NAG	H	1	15/15	0.98	0.05	126,130,134,134	0
4	SIA	F	3	20/21	0.99	0.04	99,104,110,110	0
4	GAL	F	2	11/12	0.99	0.03	82,86,89,91	0
4	GAL	H	2	11/12	0.99	0.03	79,82,87,88	0

*Continued on next page...*

*Continued from previous page...*

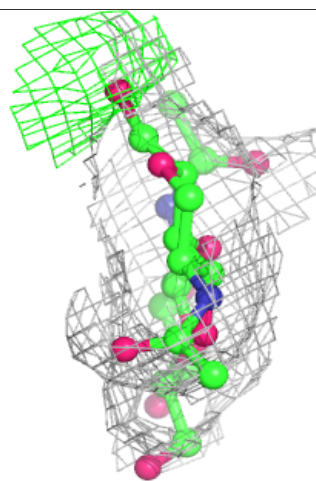
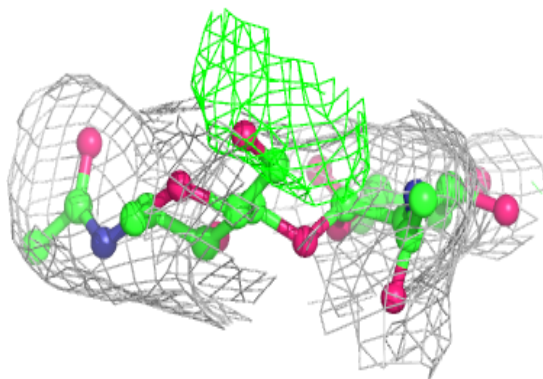
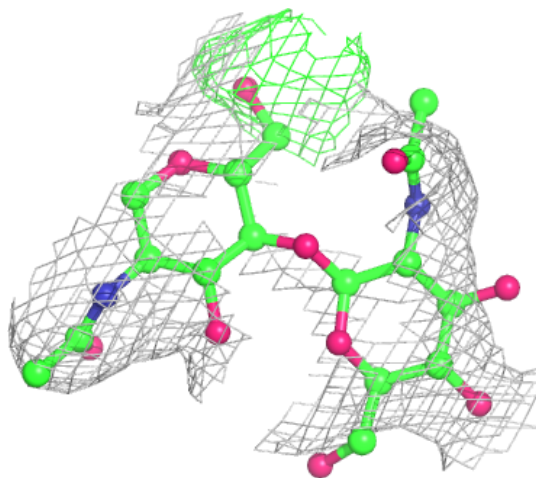
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SIA	H	3	20/21	1.00	0.04	100,106,113,114	0

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:**

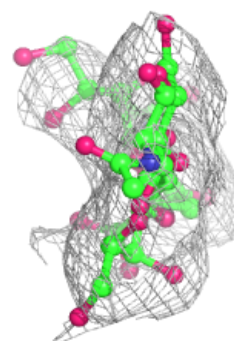
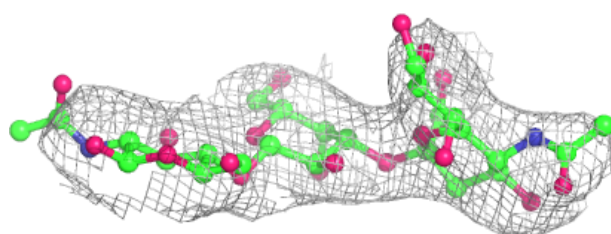
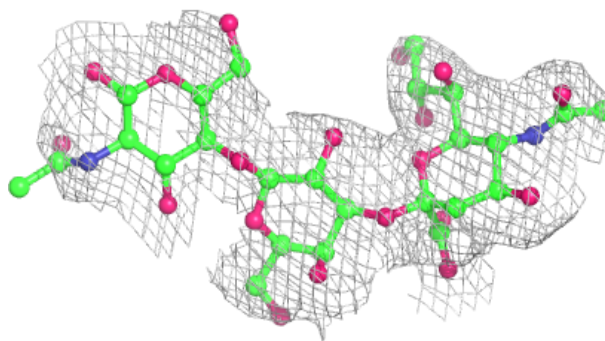
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



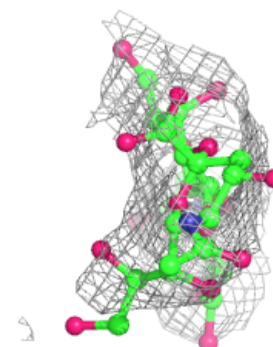
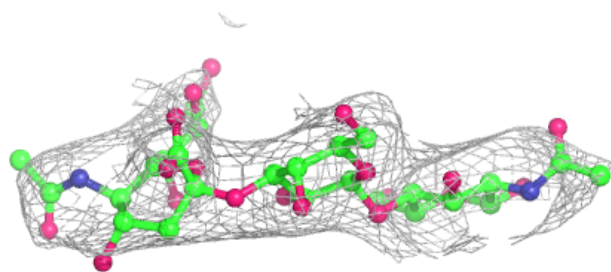
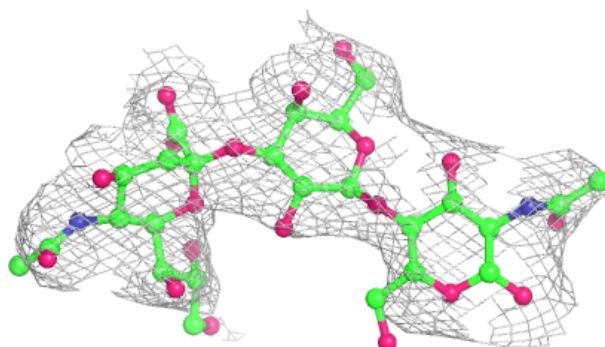


**Electron density around Chain F:**

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

**Electron density around Chain H:**

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



## 6.4 Ligands [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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
5	NAG	A	401	14/15	0.94	0.06	157,162,164,166	0
5	NAG	C	601	14/15	0.96	0.06	139,143,145,145	0
5	NAG	B	501	14/15	0.97	0.08	148,154,157,158	0
5	NAG	D	501	14/15	0.97	0.06	143,149,154,154	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.