



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 11, 2023 – 03:53 PM EDT

PDB ID : 7RDH  
Title : Crystal structure of the de novo designed binding protein H3mb in complex with the 1968 influenza A virus hemagglutinin  
Authors : Kadam, R.U.; Wilson, I.A.  
Deposited on : 2021-07-09  
Resolution : 2.75 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 : 2.35.1  
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) : 2.35.1

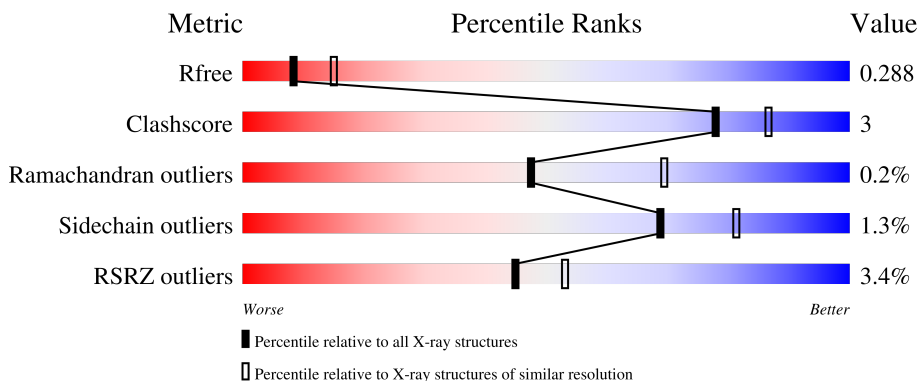
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	323	 4% 89% 9%
1	C	323	 2% 91% 7%
1	E	323	 3% 88% 10%
2	B	239	 3% 65% 6% 28%
2	D	239	 2% 72% 6% 23%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	239	<p>3% 65% 7% 28%</p>
3	G	77	<p>1% 68% 5% 26%</p>
3	H	77	<p>9% 68% 6% 26%</p>
4	I	2	<p>100%</p>
4	J	2	<p>50% 50%</p>
4	K	2	<p>100%</p>
4	M	2	<p>100%</p>
4	O	2	<p>100%</p>
5	L	3	<p>67% 33%</p>
5	N	3	<p>100%</p>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 12887 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Hemagglutinin HA1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	316	Total 2463	C 1545	N 431	O 473	S 14	0	6	0
1	C	316	Total 2461	C 1543	N 432	O 472	S 14	0	4	0
1	E	317	Total 2450	C 1532	N 431	O 473	S 14	0	3	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	expression tag	UNP Q91MA7
A	8	ASP	-	expression tag	UNP Q91MA7
A	9	PRO	-	expression tag	UNP Q91MA7
A	10	GLY	-	expression tag	UNP Q91MA7
C	7	ALA	-	expression tag	UNP Q91MA7
C	8	ASP	-	expression tag	UNP Q91MA7
C	9	PRO	-	expression tag	UNP Q91MA7
C	10	GLY	-	expression tag	UNP Q91MA7
E	7	ALA	-	expression tag	UNP Q91MA7
E	8	ASP	-	expression tag	UNP Q91MA7
E	9	PRO	-	expression tag	UNP Q91MA7
E	10	GLY	-	expression tag	UNP Q91MA7

- Molecule 2 is a protein called Hemagglutinin HA2 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	172	Total 1422	C 885	N 248	O 282	S 7	0	5	0
2	D	185	Total 1495	C 927	N 261	O 300	S 7	0	3	0
2	F	171	Total 1427	C 889	N 247	O 285	S 6	0	6	0

There are 192 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	123	GLY	ARG	engineered mutation	UNP Q91MA7
B	177	SER	-	expression tag	UNP Q91MA7
B	178	GLY	-	expression tag	UNP Q91MA7
B	179	GLY	-	expression tag	UNP Q91MA7
B	180	GLY	-	expression tag	UNP Q91MA7
B	181	GLY	-	expression tag	UNP Q91MA7
B	182	LEU	-	expression tag	UNP Q91MA7
B	183	ASN	-	expression tag	UNP Q91MA7
B	184	ASP	-	expression tag	UNP Q91MA7
B	185	ILE	-	expression tag	UNP Q91MA7
B	186	PHE	-	expression tag	UNP Q91MA7
B	187	GLU	-	expression tag	UNP Q91MA7
B	188	ALA	-	expression tag	UNP Q91MA7
B	189	GLN	-	expression tag	UNP Q91MA7
B	190	LYS	-	expression tag	UNP Q91MA7
B	191	ILE	-	expression tag	UNP Q91MA7
B	192	GLU	-	expression tag	UNP Q91MA7
B	193	TRP	-	expression tag	UNP Q91MA7
B	194	HIS	-	expression tag	UNP Q91MA7
B	195	GLU	-	expression tag	UNP Q91MA7
B	196	ARG	-	expression tag	UNP Q91MA7
B	197	LEU	-	expression tag	UNP Q91MA7
B	198	VAL	-	expression tag	UNP Q91MA7
B	199	PRO	-	expression tag	UNP Q91MA7
B	200	ARG	-	expression tag	UNP Q91MA7
B	201	GLY	-	expression tag	UNP Q91MA7
B	202	SER	-	expression tag	UNP Q91MA7
B	203	PRO	-	expression tag	UNP Q91MA7
B	204	GLY	-	expression tag	UNP Q91MA7
B	205	SER	-	expression tag	UNP Q91MA7
B	206	GLY	-	expression tag	UNP Q91MA7
B	207	TYR	-	expression tag	UNP Q91MA7
B	208	ILE	-	expression tag	UNP Q91MA7
B	209	PRO	-	expression tag	UNP Q91MA7
B	210	GLU	-	expression tag	UNP Q91MA7
B	211	ALA	-	expression tag	UNP Q91MA7
B	212	PRO	-	expression tag	UNP Q91MA7
B	213	ARG	-	expression tag	UNP Q91MA7
B	214	ASP	-	expression tag	UNP Q91MA7
B	215	GLY	-	expression tag	UNP Q91MA7
B	216	GLN	-	expression tag	UNP Q91MA7
B	217	ALA	-	expression tag	UNP Q91MA7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	218	TYR	-	expression tag	UNP Q91MA7
B	219	VAL	-	expression tag	UNP Q91MA7
B	220	ARG	-	expression tag	UNP Q91MA7
B	221	LYS	-	expression tag	UNP Q91MA7
B	222	ASP	-	expression tag	UNP Q91MA7
B	223	GLY	-	expression tag	UNP Q91MA7
B	224	GLU	-	expression tag	UNP Q91MA7
B	225	TRP	-	expression tag	UNP Q91MA7
B	226	VAL	-	expression tag	UNP Q91MA7
B	227	LEU	-	expression tag	UNP Q91MA7
B	228	LEU	-	expression tag	UNP Q91MA7
B	229	SER	-	expression tag	UNP Q91MA7
B	230	THR	-	expression tag	UNP Q91MA7
B	231	PHE	-	expression tag	UNP Q91MA7
B	232	LEU	-	expression tag	UNP Q91MA7
B	233	GLY	-	expression tag	UNP Q91MA7
B	234	HIS	-	expression tag	UNP Q91MA7
B	235	HIS	-	expression tag	UNP Q91MA7
B	236	HIS	-	expression tag	UNP Q91MA7
B	237	HIS	-	expression tag	UNP Q91MA7
B	238	HIS	-	expression tag	UNP Q91MA7
B	239	HIS	-	expression tag	UNP Q91MA7
D	123	GLY	ARG	engineered mutation	UNP Q91MA7
D	177	SER	-	expression tag	UNP Q91MA7
D	178	GLY	-	expression tag	UNP Q91MA7
D	179	GLY	-	expression tag	UNP Q91MA7
D	180	GLY	-	expression tag	UNP Q91MA7
D	181	GLY	-	expression tag	UNP Q91MA7
D	182	LEU	-	expression tag	UNP Q91MA7
D	183	ASN	-	expression tag	UNP Q91MA7
D	184	ASP	-	expression tag	UNP Q91MA7
D	185	ILE	-	expression tag	UNP Q91MA7
D	186	PHE	-	expression tag	UNP Q91MA7
D	187	GLU	-	expression tag	UNP Q91MA7
D	188	ALA	-	expression tag	UNP Q91MA7
D	189	GLN	-	expression tag	UNP Q91MA7
D	190	LYS	-	expression tag	UNP Q91MA7
D	191	ILE	-	expression tag	UNP Q91MA7
D	192	GLU	-	expression tag	UNP Q91MA7
D	193	TRP	-	expression tag	UNP Q91MA7
D	194	HIS	-	expression tag	UNP Q91MA7
D	195	GLU	-	expression tag	UNP Q91MA7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	196	ARG	-	expression tag	UNP Q91MA7
D	197	LEU	-	expression tag	UNP Q91MA7
D	198	VAL	-	expression tag	UNP Q91MA7
D	199	PRO	-	expression tag	UNP Q91MA7
D	200	ARG	-	expression tag	UNP Q91MA7
D	201	GLY	-	expression tag	UNP Q91MA7
D	202	SER	-	expression tag	UNP Q91MA7
D	203	PRO	-	expression tag	UNP Q91MA7
D	204	GLY	-	expression tag	UNP Q91MA7
D	205	SER	-	expression tag	UNP Q91MA7
D	206	GLY	-	expression tag	UNP Q91MA7
D	207	TYR	-	expression tag	UNP Q91MA7
D	208	ILE	-	expression tag	UNP Q91MA7
D	209	PRO	-	expression tag	UNP Q91MA7
D	210	GLU	-	expression tag	UNP Q91MA7
D	211	ALA	-	expression tag	UNP Q91MA7
D	212	PRO	-	expression tag	UNP Q91MA7
D	213	ARG	-	expression tag	UNP Q91MA7
D	214	ASP	-	expression tag	UNP Q91MA7
D	215	GLY	-	expression tag	UNP Q91MA7
D	216	GLN	-	expression tag	UNP Q91MA7
D	217	ALA	-	expression tag	UNP Q91MA7
D	218	TYR	-	expression tag	UNP Q91MA7
D	219	VAL	-	expression tag	UNP Q91MA7
D	220	ARG	-	expression tag	UNP Q91MA7
D	221	LYS	-	expression tag	UNP Q91MA7
D	222	ASP	-	expression tag	UNP Q91MA7
D	223	GLY	-	expression tag	UNP Q91MA7
D	224	GLU	-	expression tag	UNP Q91MA7
D	225	TRP	-	expression tag	UNP Q91MA7
D	226	VAL	-	expression tag	UNP Q91MA7
D	227	LEU	-	expression tag	UNP Q91MA7
D	228	LEU	-	expression tag	UNP Q91MA7
D	229	SER	-	expression tag	UNP Q91MA7
D	230	THR	-	expression tag	UNP Q91MA7
D	231	PHE	-	expression tag	UNP Q91MA7
D	232	LEU	-	expression tag	UNP Q91MA7
D	233	GLY	-	expression tag	UNP Q91MA7
D	234	HIS	-	expression tag	UNP Q91MA7
D	235	HIS	-	expression tag	UNP Q91MA7
D	236	HIS	-	expression tag	UNP Q91MA7
D	237	HIS	-	expression tag	UNP Q91MA7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	238	HIS	-	expression tag	UNP Q91MA7
D	239	HIS	-	expression tag	UNP Q91MA7
F	123	GLY	ARG	engineered mutation	UNP Q91MA7
F	177	SER	-	expression tag	UNP Q91MA7
F	178	GLY	-	expression tag	UNP Q91MA7
F	179	GLY	-	expression tag	UNP Q91MA7
F	180	GLY	-	expression tag	UNP Q91MA7
F	181	GLY	-	expression tag	UNP Q91MA7
F	182	LEU	-	expression tag	UNP Q91MA7
F	183	ASN	-	expression tag	UNP Q91MA7
F	184	ASP	-	expression tag	UNP Q91MA7
F	185	ILE	-	expression tag	UNP Q91MA7
F	186	PHE	-	expression tag	UNP Q91MA7
F	187	GLU	-	expression tag	UNP Q91MA7
F	188	ALA	-	expression tag	UNP Q91MA7
F	189	GLN	-	expression tag	UNP Q91MA7
F	190	LYS	-	expression tag	UNP Q91MA7
F	191	ILE	-	expression tag	UNP Q91MA7
F	192	GLU	-	expression tag	UNP Q91MA7
F	193	TRP	-	expression tag	UNP Q91MA7
F	194	HIS	-	expression tag	UNP Q91MA7
F	195	GLU	-	expression tag	UNP Q91MA7
F	196	ARG	-	expression tag	UNP Q91MA7
F	197	LEU	-	expression tag	UNP Q91MA7
F	198	VAL	-	expression tag	UNP Q91MA7
F	199	PRO	-	expression tag	UNP Q91MA7
F	200	ARG	-	expression tag	UNP Q91MA7
F	201	GLY	-	expression tag	UNP Q91MA7
F	202	SER	-	expression tag	UNP Q91MA7
F	203	PRO	-	expression tag	UNP Q91MA7
F	204	GLY	-	expression tag	UNP Q91MA7
F	205	SER	-	expression tag	UNP Q91MA7
F	206	GLY	-	expression tag	UNP Q91MA7
F	207	TYR	-	expression tag	UNP Q91MA7
F	208	ILE	-	expression tag	UNP Q91MA7
F	209	PRO	-	expression tag	UNP Q91MA7
F	210	GLU	-	expression tag	UNP Q91MA7
F	211	ALA	-	expression tag	UNP Q91MA7
F	212	PRO	-	expression tag	UNP Q91MA7
F	213	ARG	-	expression tag	UNP Q91MA7
F	214	ASP	-	expression tag	UNP Q91MA7
F	215	GLY	-	expression tag	UNP Q91MA7

*Continued on next page...*



Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	216	GLN	-	expression tag	UNP Q91MA7
F	217	ALA	-	expression tag	UNP Q91MA7
F	218	TYR	-	expression tag	UNP Q91MA7
F	219	VAL	-	expression tag	UNP Q91MA7
F	220	ARG	-	expression tag	UNP Q91MA7
F	221	LYS	-	expression tag	UNP Q91MA7
F	222	ASP	-	expression tag	UNP Q91MA7
F	223	GLY	-	expression tag	UNP Q91MA7
F	224	GLU	-	expression tag	UNP Q91MA7
F	225	TRP	-	expression tag	UNP Q91MA7
F	226	VAL	-	expression tag	UNP Q91MA7
F	227	LEU	-	expression tag	UNP Q91MA7
F	228	LEU	-	expression tag	UNP Q91MA7
F	229	SER	-	expression tag	UNP Q91MA7
F	230	THR	-	expression tag	UNP Q91MA7
F	231	PHE	-	expression tag	UNP Q91MA7
F	232	LEU	-	expression tag	UNP Q91MA7
F	233	GLY	-	expression tag	UNP Q91MA7
F	234	HIS	-	expression tag	UNP Q91MA7
F	235	HIS	-	expression tag	UNP Q91MA7
F	236	HIS	-	expression tag	UNP Q91MA7
F	237	HIS	-	expression tag	UNP Q91MA7
F	238	HIS	-	expression tag	UNP Q91MA7
F	239	HIS	-	expression tag	UNP Q91MA7

- Molecule 3 is a protein called De novo designed protein H3mb.

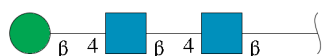
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	57	Total	C	N	O	S	0	0	0
			412	267	78	66	1			
3	H	57	Total	C	N	O	S	0	0	0
			375	242	71	61	1			

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



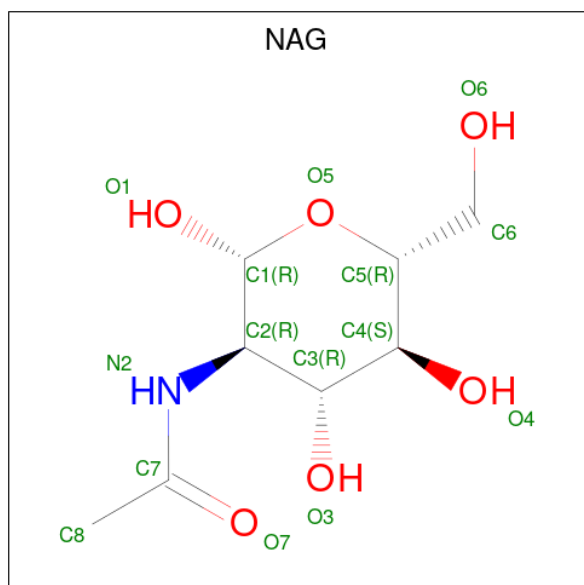
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	I	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	J	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	K	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	M	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	O	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 5 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



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

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



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

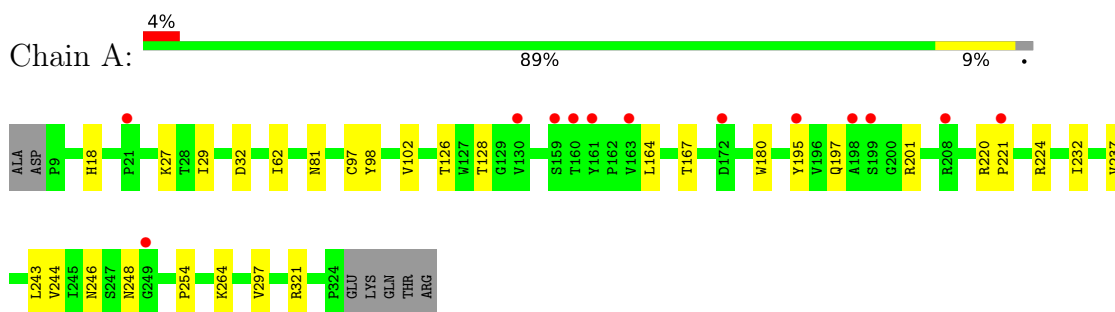
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	20	Total	O	0	0
			20	20		
7	B	12	Total	O	0	0
			12	12		
7	C	26	Total	O	0	0
			26	26		
7	D	15	Total	O	0	0
			15	15		
7	E	10	Total	O	0	0
			10	10		
7	F	5	Total	O	0	0
			5	5		
7	G	3	Total	O	0	0
			3	3		
7	H	3	Total	O	0	0
			3	3		

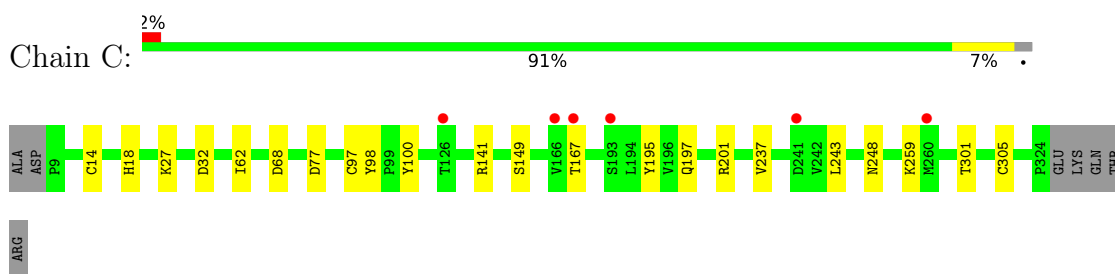
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

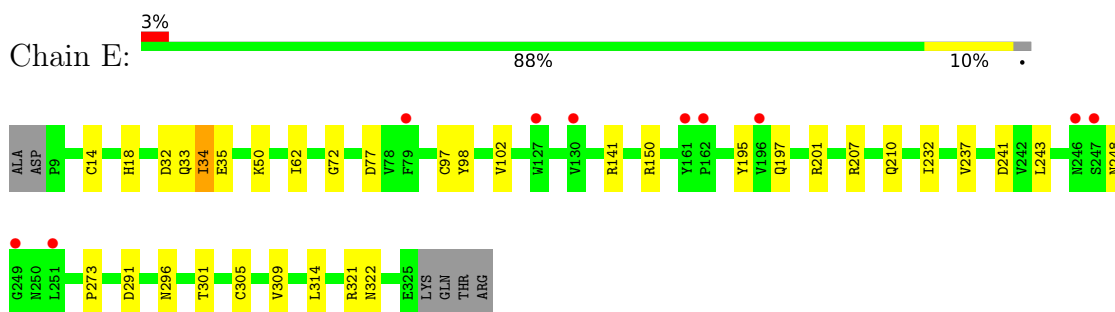
- Molecule 1: Hemagglutinin HA1 chain



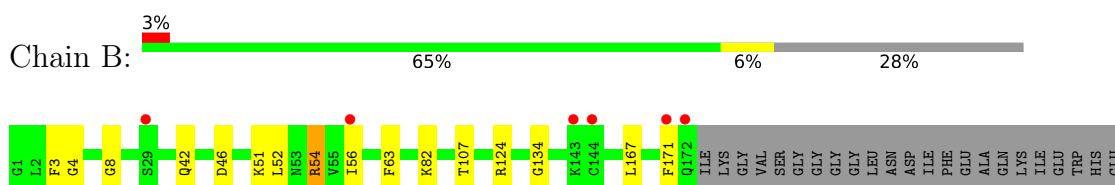
- Molecule 1: Hemagglutinin HA1 chain



- Molecule 1: Hemagglutinin HA1 chain



- Molecule 2: Hemagglutinin HA2 chain





MAG1  
MAG2

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

Chain K:  100%MAG1  
MAG2

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

Chain M:  100%MAG1  
MAG2

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

Chain O:  100%MAG1  
MAG2

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

Chain L:  67% 33%MAG1  
MAG2  
EMAS

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

Chain N:  100%MAG1  
MAG2  
EMAS

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.92Å 240.84Å 70.70Å 90.00° 117.30° 90.00°	Depositor
Resolution (Å)	43.24 – 2.75 43.24 – 2.75	Depositor EDS
% Data completeness (in resolution range)	89.0 (43.24-2.75) 89.0 (43.24-2.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.29 (at 2.77Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.232 , 0.289 0.232 , 0.288	Depositor DCC
$R_{free}$ test set	2359 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	57.3	Xtrriage
Anisotropy	0.725	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 44.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.036 for l,-k,h	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12887	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.87% 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: 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	0.25	0/2523	0.49	0/3439
1	C	0.24	0/2518	0.48	0/3432
1	E	0.25	0/2505	0.49	0/3412
2	B	0.26	0/1446	0.47	0/1943
2	D	0.26	0/1519	0.46	0/2040
2	F	0.25	0/1454	0.46	0/1954
3	G	0.26	0/418	0.43	0/564
3	H	0.25	0/381	0.43	0/520
All	All	0.25	0/12764	0.48	0/17304

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	2463	0	2406	17	0
1	C	2461	0	2407	11	0
1	E	2450	0	2395	19	0
2	B	1422	0	1340	14	0
2	D	1495	0	1414	11	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	1427	0	1351	9	0
3	G	412	0	378	4	0
3	H	375	0	297	1	0
4	I	28	0	25	0	0
4	J	28	0	25	1	0
4	K	28	0	25	0	0
4	M	28	0	25	0	0
4	O	28	0	25	0	0
5	L	39	0	34	0	0
5	N	39	0	34	0	0
6	A	14	0	13	0	0
6	C	28	0	26	0	0
6	E	28	0	26	1	0
7	A	20	0	0	2	0
7	B	12	0	0	1	0
7	C	26	0	0	1	0
7	D	15	0	0	1	0
7	E	10	0	0	0	0
7	F	5	0	0	0	0
7	G	3	0	0	0	0
7	H	3	0	0	0	0
All	All	12887	0	12246	72	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 72 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:150:ARG:HH12	6:E:401:NAG:H81	1.57	0.70
1:A:167:THR:HG22	1:A:244:VAL:HG22	1.79	0.64
1:E:102:VAL:HG22	1:E:232:ILE:HB	1.82	0.62
1:C:97:CYS:SG	1:C:98:TYR:N	2.73	0.61
1:C:77:ASP:OD2	1:C:141:ARG:NH1	2.32	0.60

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	320/323 (99%)	308 (96%)	11 (3%)	1 (0%)	41	60
1	C	318/323 (98%)	307 (96%)	10 (3%)	1 (0%)	41	60
1	E	318/323 (98%)	305 (96%)	12 (4%)	1 (0%)	41	60
2	B	175/239 (73%)	167 (95%)	8 (5%)	0	100	100
2	D	186/239 (78%)	174 (94%)	12 (6%)	0	100	100
2	F	175/239 (73%)	165 (94%)	10 (6%)	0	100	100
3	G	55/77 (71%)	50 (91%)	5 (9%)	0	100	100
3	H	55/77 (71%)	50 (91%)	5 (9%)	0	100	100
All	All	1602/1840 (87%)	1526 (95%)	73 (5%)	3 (0%)	47	69

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	62	ILE
1	C	62	ILE
1	E	62	ILE

### 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	280/283 (99%)	278 (99%)	2 (1%)	84	89
1	C	280/283 (99%)	278 (99%)	2 (1%)	84	89

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	279/283 (99%)	275 (99%)	4 (1%)	67	79
2	B	149/199 (75%)	148 (99%)	1 (1%)	84	89
2	D	157/199 (79%)	156 (99%)	1 (1%)	86	90
2	F	151/199 (76%)	148 (98%)	3 (2%)	55	72
3	G	32/70 (46%)	31 (97%)	1 (3%)	40	60
3	H	22/70 (31%)	19 (86%)	3 (14%)	3	5
All	All	1350/1586 (85%)	1333 (99%)	17 (1%)	69	81

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	H	9	TRP
3	H	42	ASP
1	E	34	ILE
1	E	291	ASP
1	E	321	ARG

Sometimes 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](#)

16 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
4	NAG	I	1	4,1	14,14,15	0.32	0	17,19,21	0.39	0
4	NAG	I	2	4	14,14,15	0.21	0	17,19,21	0.38	0
4	NAG	J	1	4,1	14,14,15	0.23	0	17,19,21	0.41	0
4	NAG	J	2	4	14,14,15	0.20	0	17,19,21	0.47	0
4	NAG	K	1	4,2	14,14,15	0.30	0	17,19,21	0.57	0
4	NAG	K	2	4	14,14,15	0.31	0	17,19,21	0.40	0
5	NAG	L	1	5,1	14,14,15	0.24	0	17,19,21	0.51	0
5	NAG	L	2	5	14,14,15	0.28	0	17,19,21	0.66	0
5	BMA	L	3	5	11,11,12	0.47	0	15,15,17	0.90	1 (6%)
4	NAG	M	1	4,1	14,14,15	0.23	0	17,19,21	0.38	0
4	NAG	M	2	4	14,14,15	0.20	0	17,19,21	0.44	0
5	NAG	N	1	5,1	14,14,15	0.34	0	17,19,21	0.41	0
5	NAG	N	2	5	14,14,15	0.24	0	17,19,21	0.39	0
5	BMA	N	3	5	11,11,12	0.63	0	15,15,17	0.73	0
4	NAG	O	1	4,1	14,14,15	0.24	0	17,19,21	0.41	0
4	NAG	O	2	4	14,14,15	0.24	0	17,19,21	0.43	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
4	NAG	I	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	I	2	4	-	2/6/23/26	0/1/1/1
4	NAG	J	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	J	2	4	-	0/6/23/26	0/1/1/1
4	NAG	K	1	4,2	-	3/6/23/26	0/1/1/1
4	NAG	K	2	4	-	0/6/23/26	0/1/1/1
5	NAG	L	1	5,1	-	0/6/23/26	0/1/1/1
5	NAG	L	2	5	-	3/6/23/26	0/1/1/1
5	BMA	L	3	5	-	0/2/19/22	0/1/1/1
4	NAG	M	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	M	2	4	-	0/6/23/26	0/1/1/1
5	NAG	N	1	5,1	-	0/6/23/26	0/1/1/1
5	NAG	N	2	5	-	0/6/23/26	0/1/1/1
5	BMA	N	3	5	-	1/2/19/22	0/1/1/1
4	NAG	O	1	4,1	-	1/6/23/26	0/1/1/1
4	NAG	O	2	4	-	2/6/23/26	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(°)
5	L	3	BMA	C1-O5-C5	2.10	115.04	112.19

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

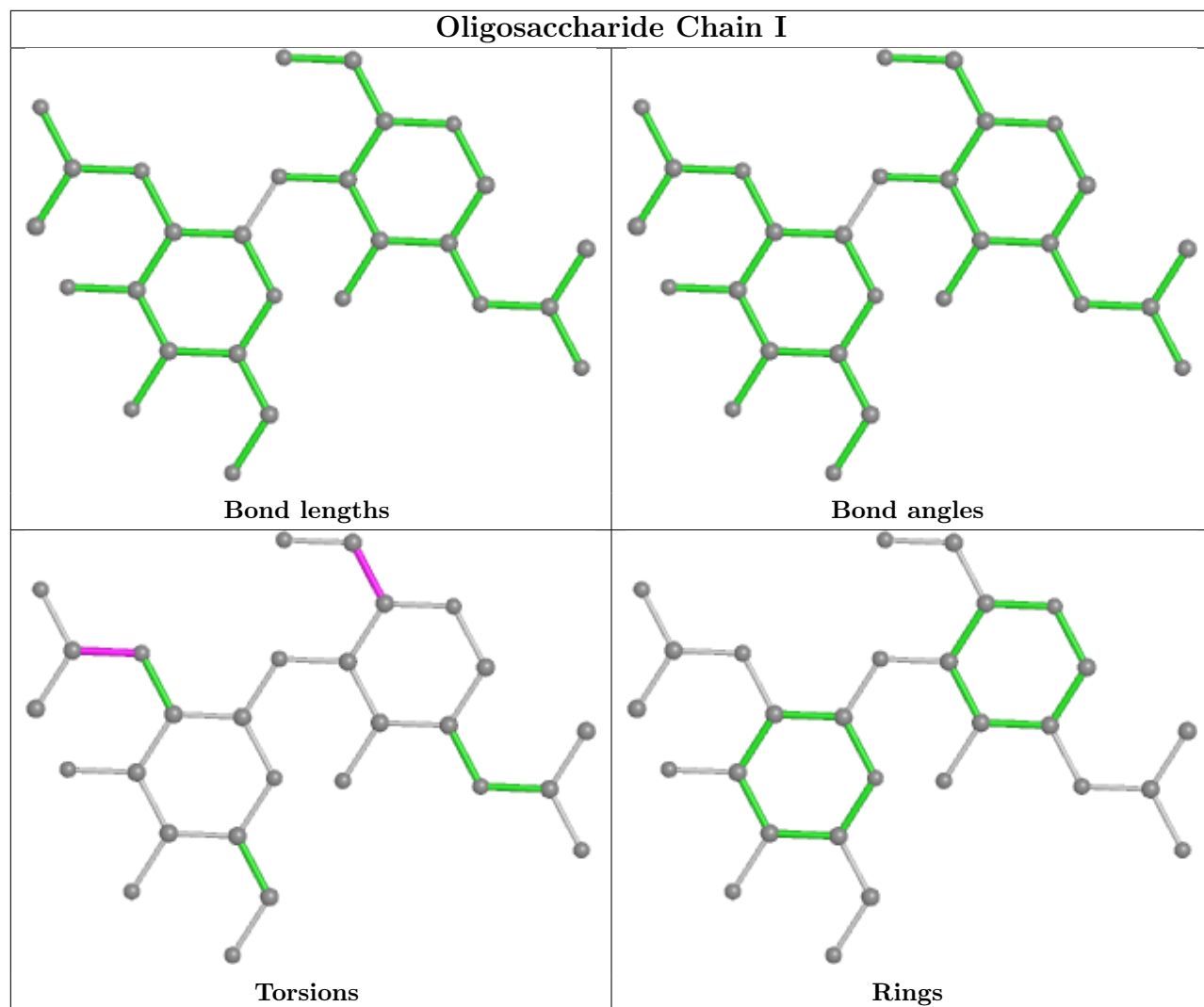
Mol	Chain	Res	Type	Atoms
4	O	2	NAG	O5-C5-C6-O6
4	K	1	NAG	O5-C5-C6-O6
4	I	1	NAG	O5-C5-C6-O6
4	K	1	NAG	C4-C5-C6-O6
4	I	2	NAG	C8-C7-N2-C2

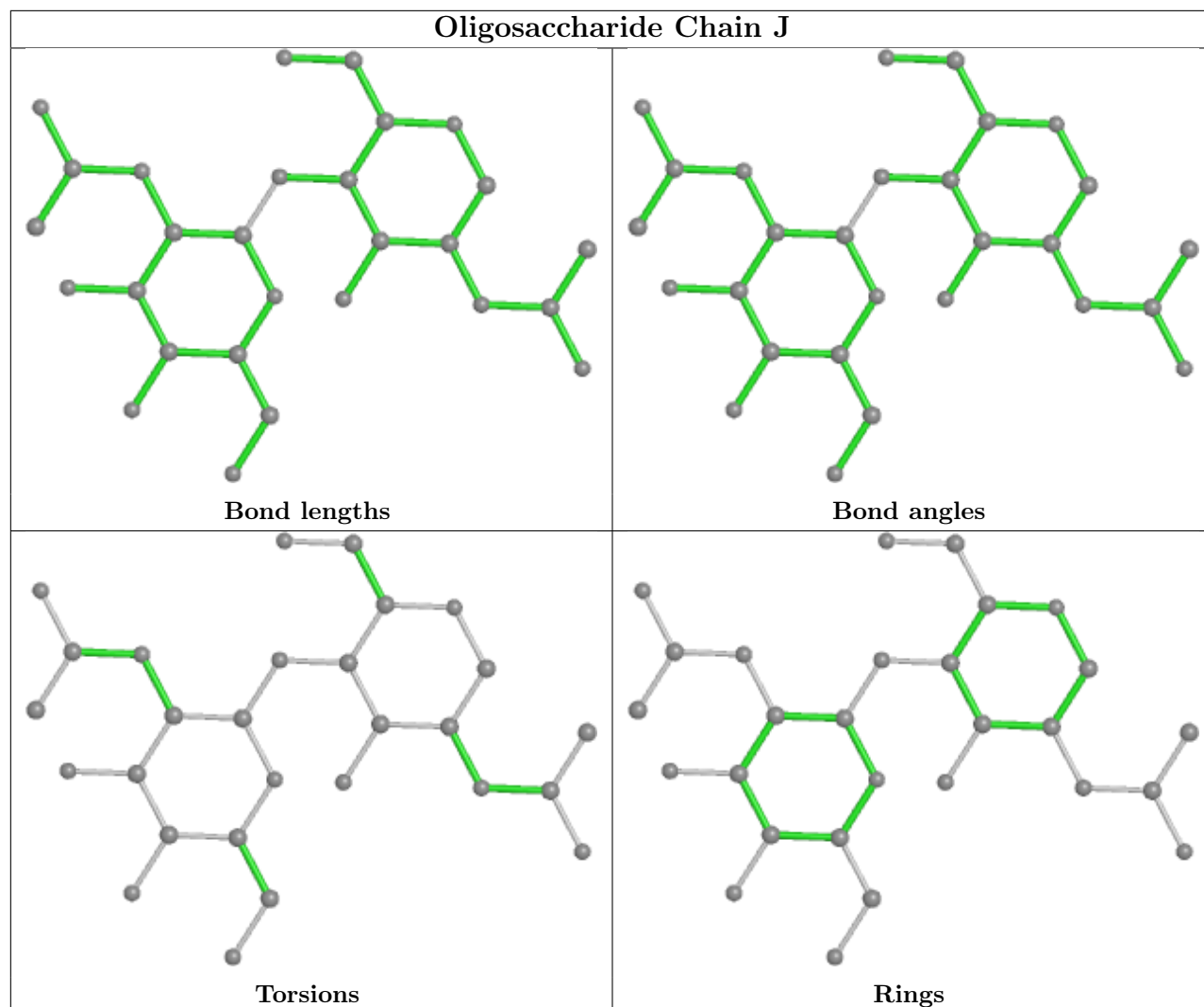
There are no ring outliers.

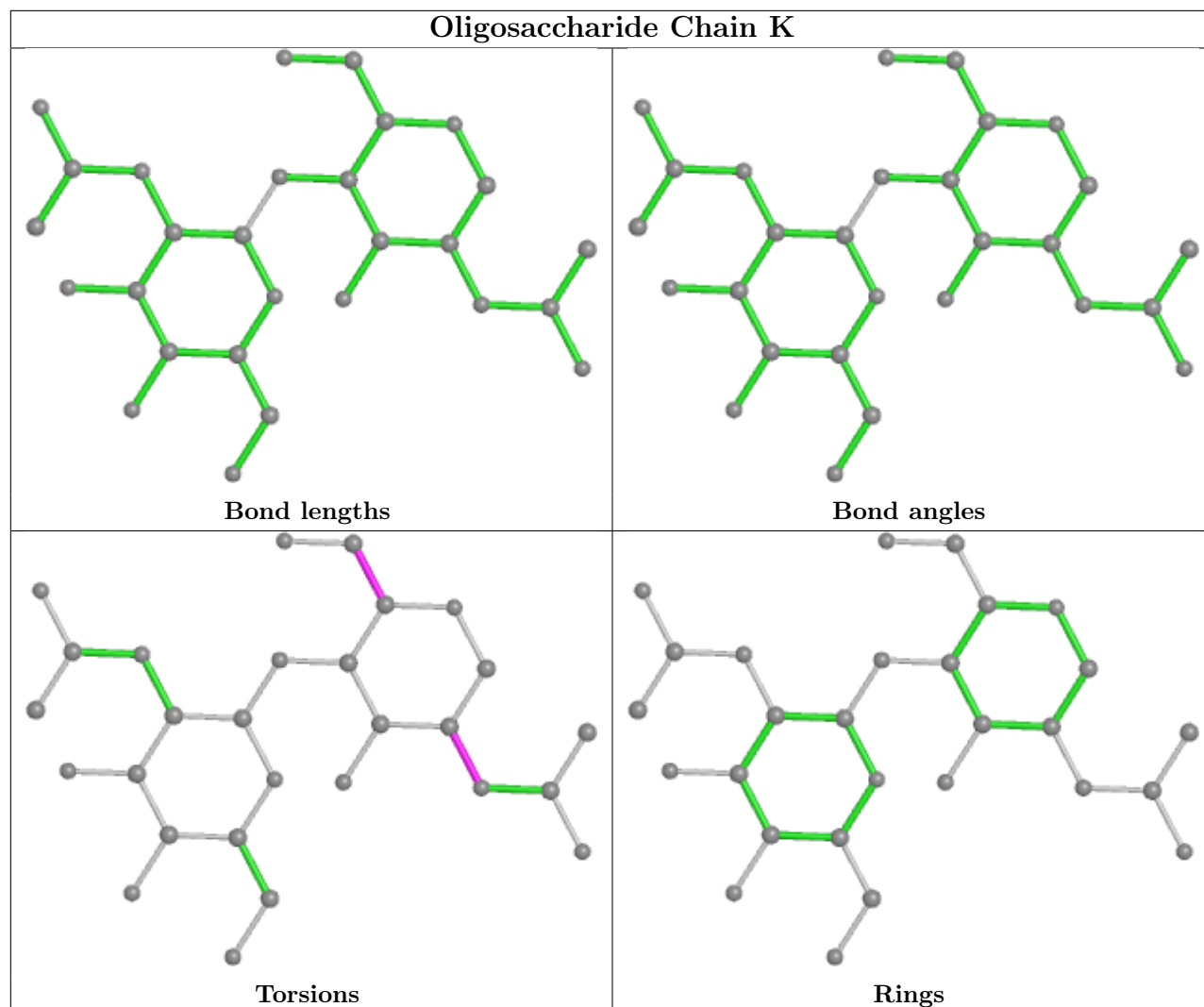
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	J	1	NAG	1	0

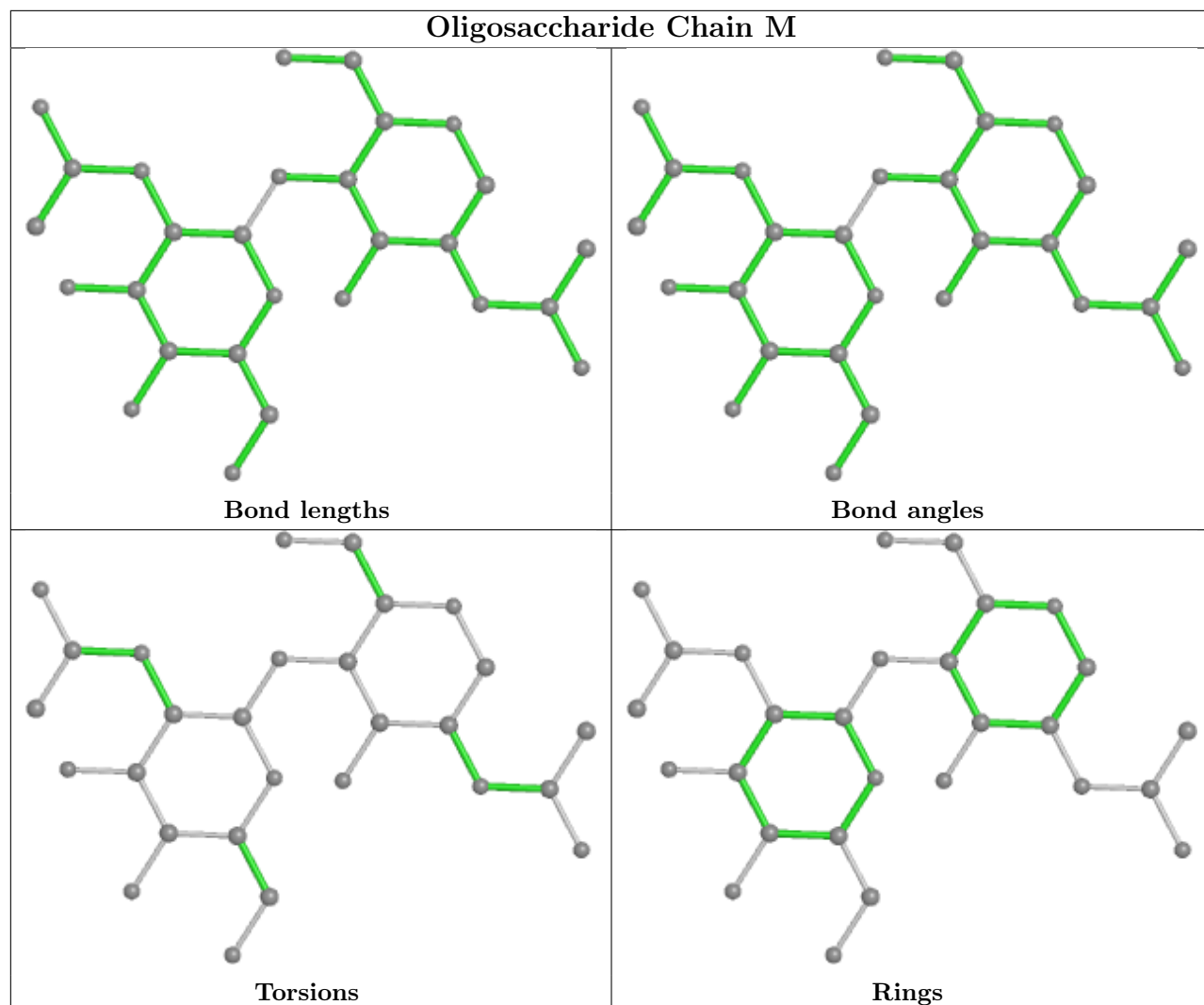
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

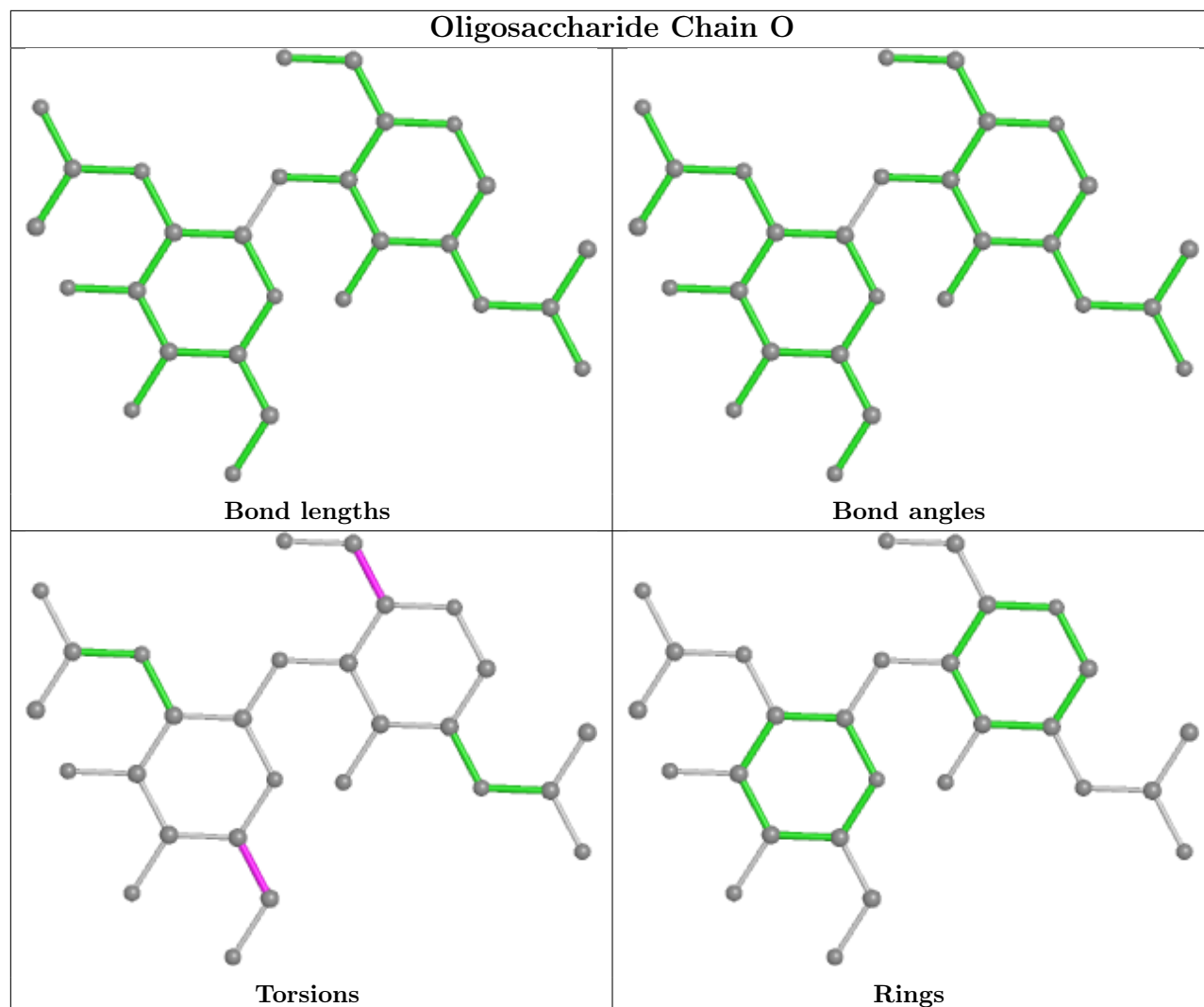


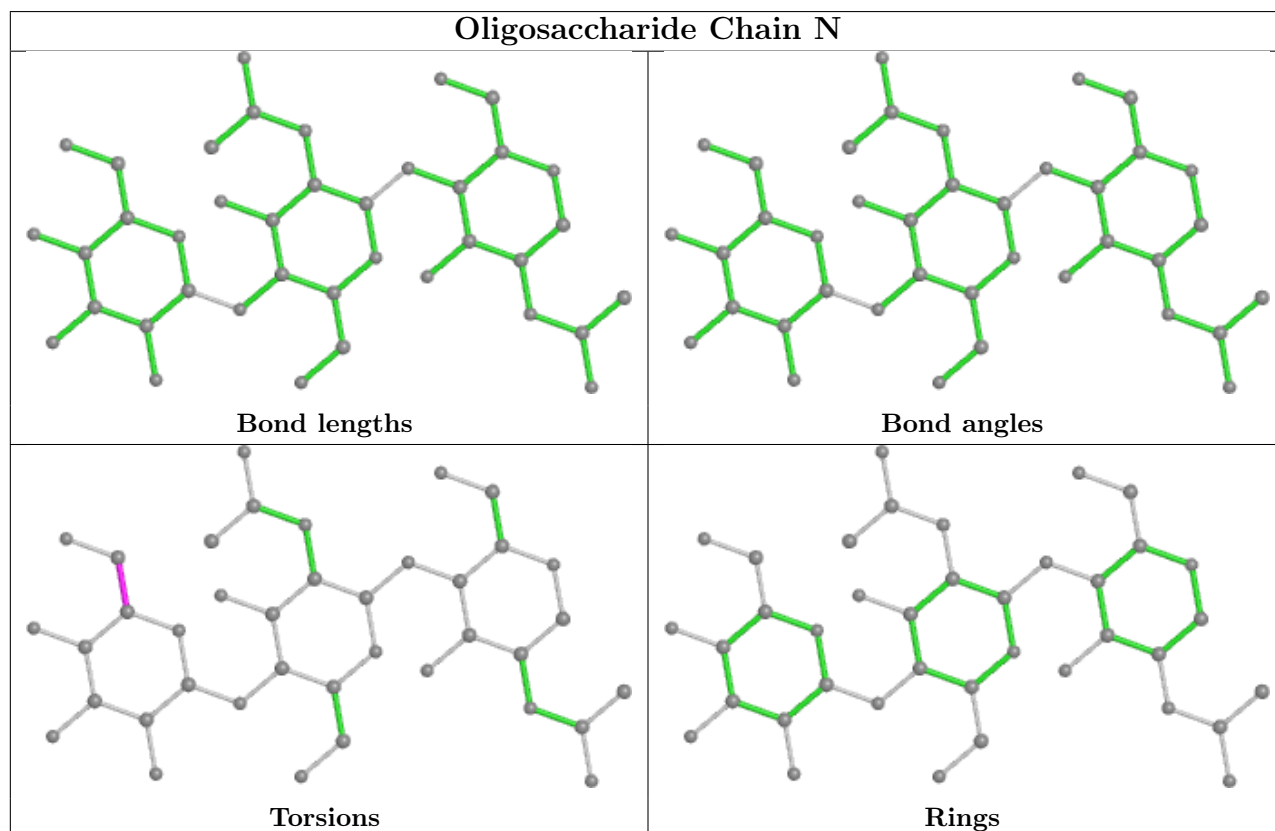
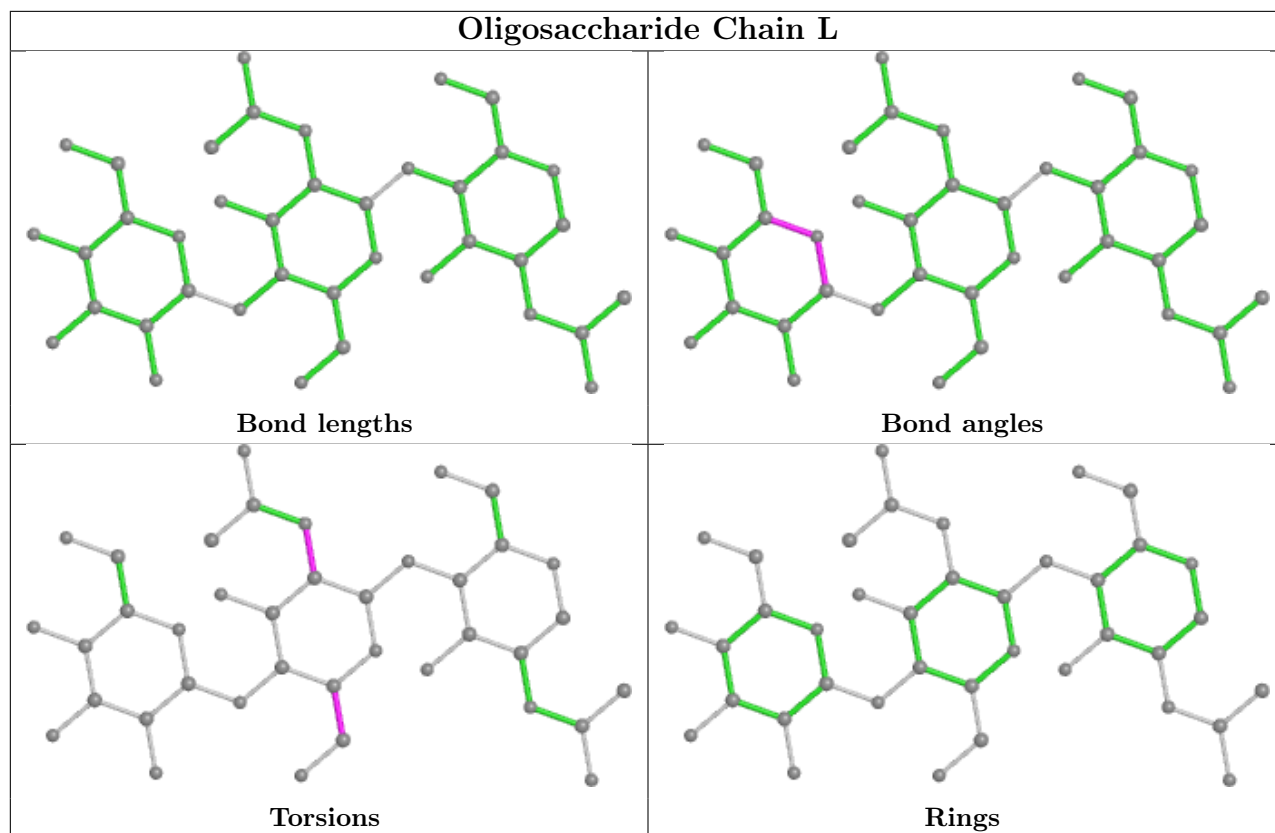












## 5.6 Ligand geometry

5 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	C	401	1	14,14,15	0.25	0	17,19,21	0.39	0
6	NAG	C	402	1	14,14,15	0.35	0	17,19,21	0.56	0
6	NAG	E	402	1	14,14,15	0.25	0	17,19,21	0.49	0
6	NAG	A	401	1	14,14,15	0.20	0	17,19,21	0.41	0
6	NAG	E	401	1	14,14,15	0.30	0	17,19,21	0.65	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	C	401	1	-	1/6/23/26	0/1/1/1
6	NAG	C	402	1	-	2/6/23/26	0/1/1/1
6	NAG	E	402	1	-	2/6/23/26	0/1/1/1
6	NAG	A	401	1	-	1/6/23/26	0/1/1/1
6	NAG	E	401	1	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	E	402	NAG	C8-C7-N2-C2
6	E	402	NAG	O7-C7-N2-C2
6	C	402	NAG	O5-C5-C6-O6
6	C	401	NAG	O5-C5-C6-O6
6	A	401	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	E	401	NAG	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

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

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

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	316/323 (97%)	0.31	13 (4%) 37 44	34, 60, 96, 113	0
1	C	316/323 (97%)	0.15	6 (1%) 66 75	33, 61, 90, 105	0
1	E	317/323 (98%)	0.26	10 (3%) 47 56	33, 70, 101, 119	0
2	B	172/239 (71%)	0.30	6 (3%) 44 52	30, 57, 80, 95	0
2	D	185/239 (77%)	0.25	5 (2%) 54 63	32, 62, 85, 97	0
2	F	171/239 (71%)	0.38	6 (3%) 44 52	34, 64, 97, 108	0
3	G	57/77 (74%)	-0.04	1 (1%) 68 76	48, 61, 84, 102	0
3	H	57/77 (74%)	0.50	7 (12%) 4 4	57, 77, 99, 116	0
All	All	1591/1840 (86%)	0.26	54 (3%) 45 53	30, 63, 96, 119	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	160	THR	4.4
1	E	196	VAL	4.2
2	F	149	ILE	4.1
1	A	172	ASP	4.1
2	B	56	ILE	4.1

### 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,

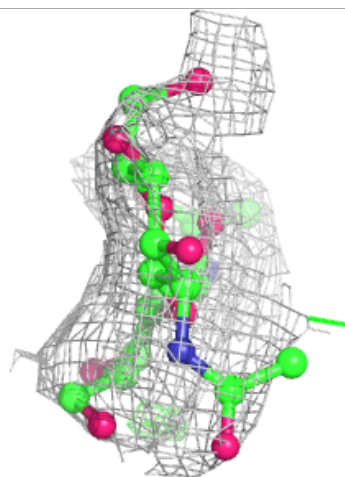
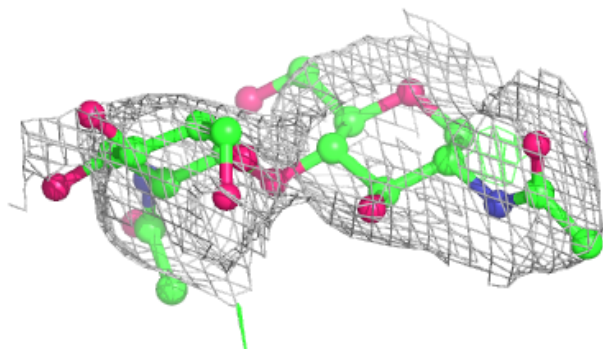
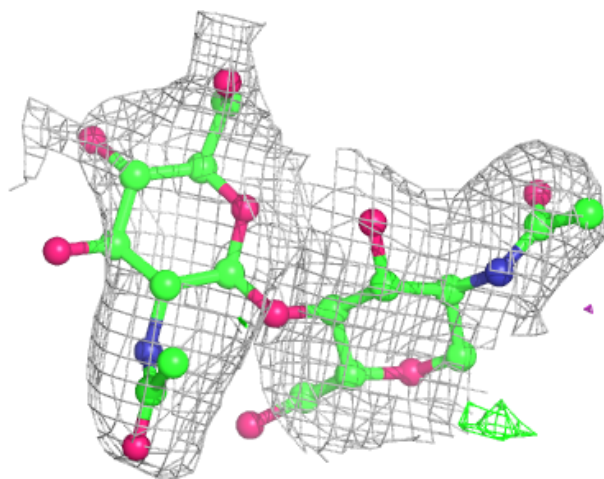
median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	K	2	14/15	0.70	0.29	89,104,122,122	0
4	NAG	J	2	14/15	0.76	0.20	76,110,119,122	0
4	NAG	I	2	14/15	0.76	0.29	91,109,128,129	0
5	BMA	N	3	11/12	0.77	0.20	110,114,131,131	0
5	BMA	L	3	11/12	0.79	0.30	109,119,129,129	0
5	NAG	L	2	14/15	0.81	0.27	87,118,124,124	0
5	NAG	N	2	14/15	0.82	0.26	73,112,120,122	0
4	NAG	K	1	14/15	0.82	0.24	76,101,112,116	0
4	NAG	O	2	14/15	0.83	0.25	103,113,131,131	0
5	NAG	L	1	14/15	0.83	0.24	52,88,99,108	0
4	NAG	I	1	14/15	0.84	0.19	68,86,125,131	0
4	NAG	M	1	14/15	0.84	0.17	60,68,84,102	0
4	NAG	M	2	14/15	0.87	0.23	67,90,105,118	0
5	NAG	N	1	14/15	0.88	0.14	41,69,85,95	0
4	NAG	O	1	14/15	0.89	0.20	77,97,107,110	0
4	NAG	J	1	14/15	0.94	0.12	51,78,97,104	0

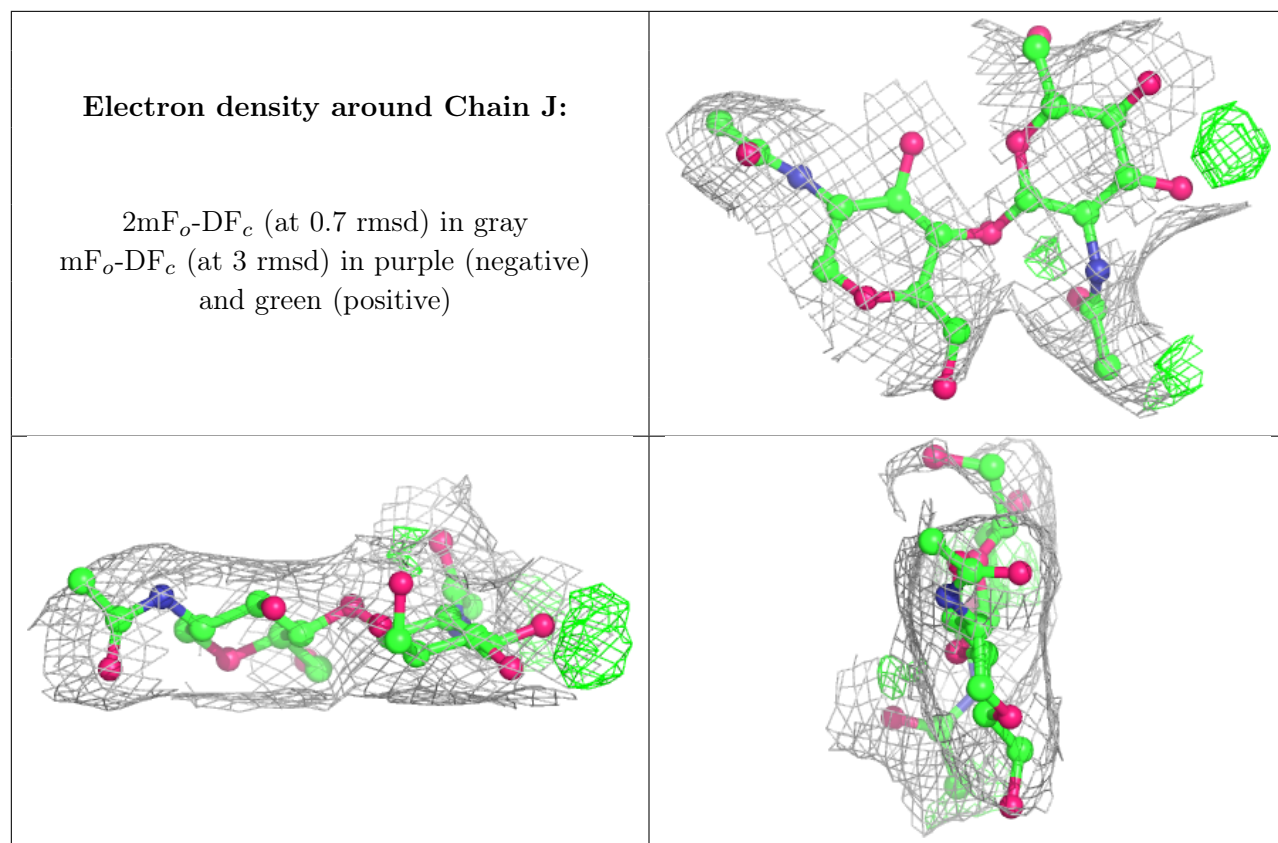
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain I:**

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

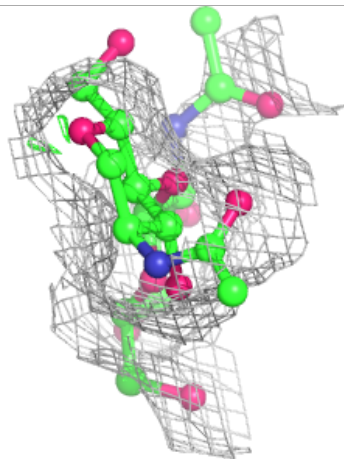
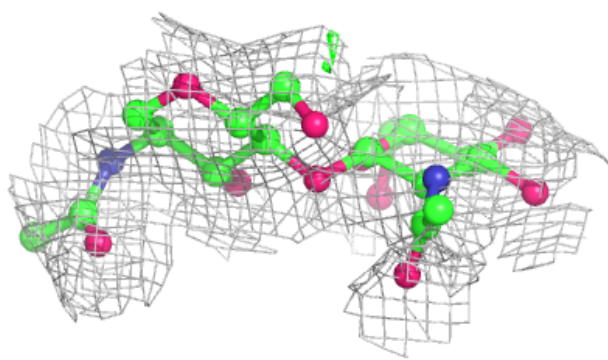
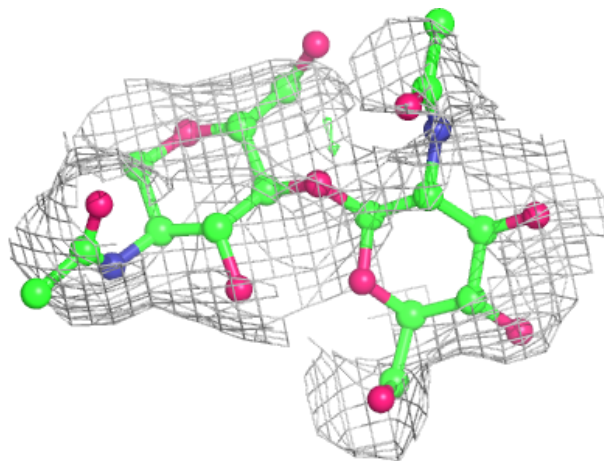






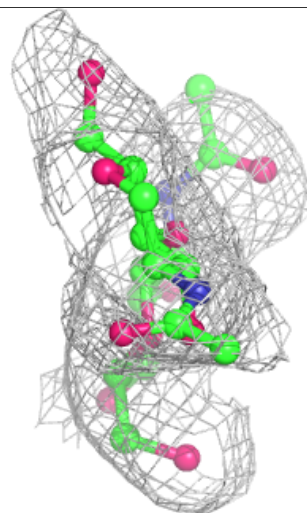
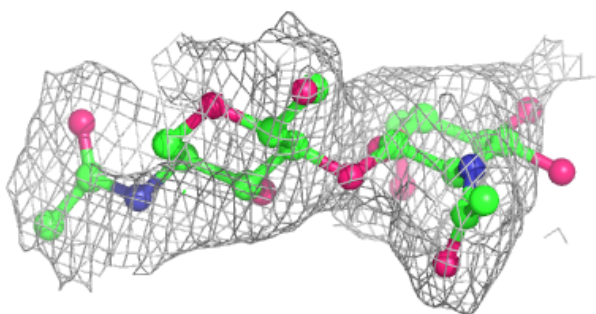
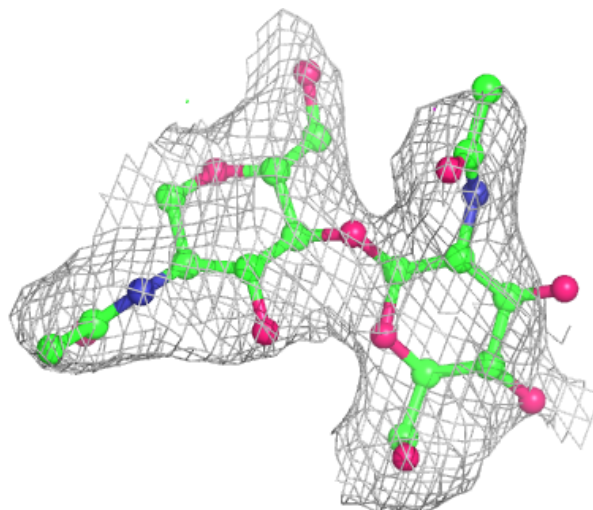
**Electron density around Chain K:**

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



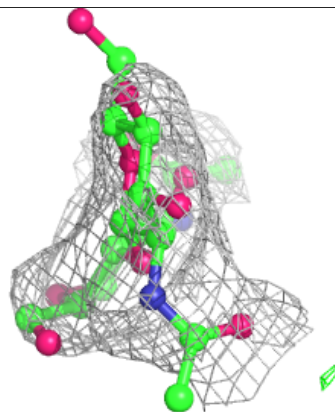
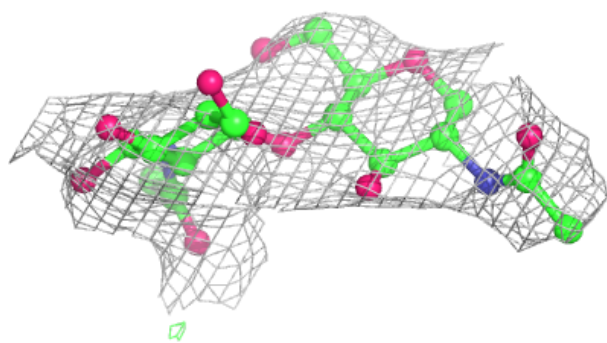
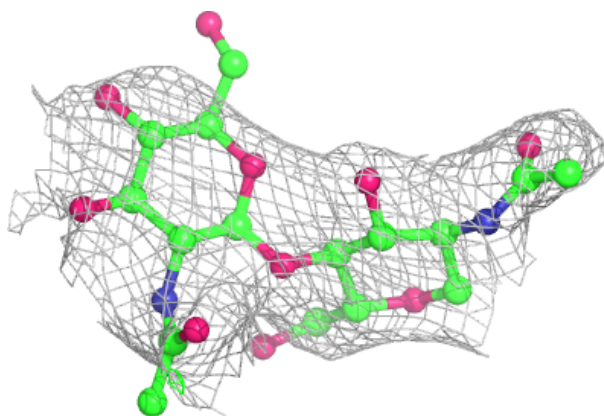
**Electron density around Chain M:**

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



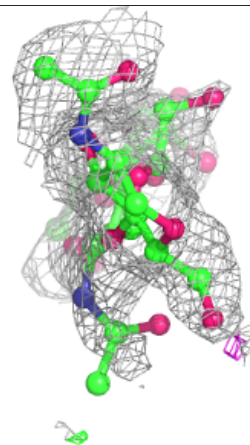
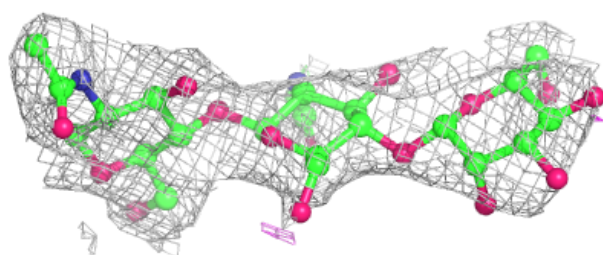
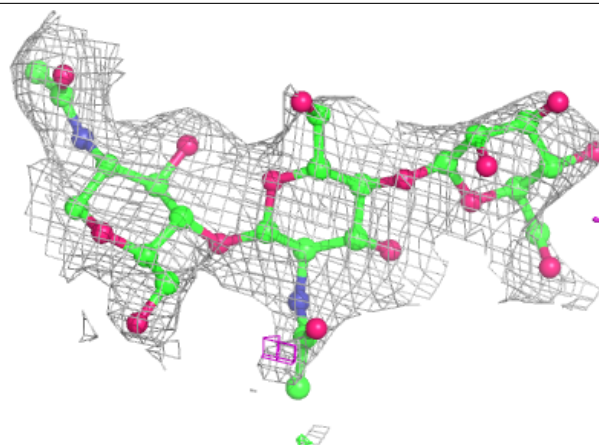
**Electron density around Chain O:**

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

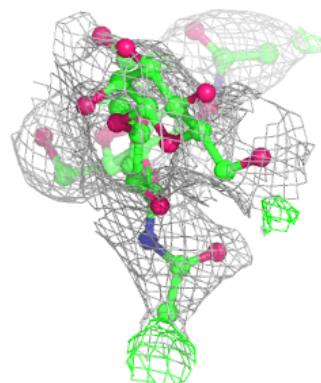
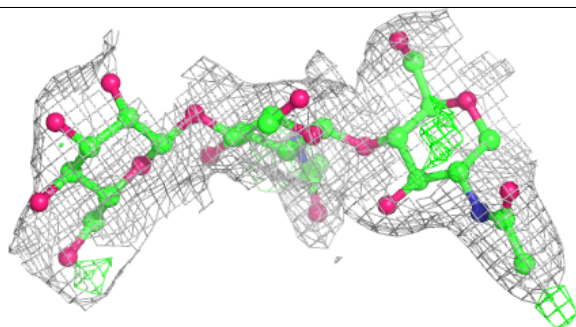
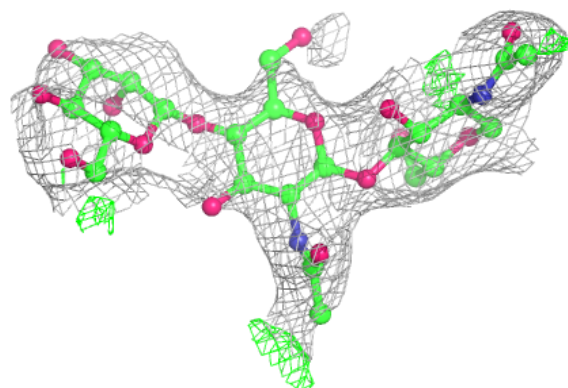


**Electron density around Chain L:**

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

**Electron density around Chain N:**

$2mF_o-DF_c$  (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
6	NAG	E	402	14/15	0.75	0.29	84,110,121,131	0
6	NAG	E	401	14/15	0.76	0.25	64,86,96,96	0
6	NAG	C	401	14/15	0.77	0.23	84,95,104,107	0
6	NAG	A	401	14/15	0.79	0.26	77,105,116,121	0
6	NAG	C	402	14/15	0.85	0.22	77,86,101,101	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.