



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 23, 2023 – 08:42 PM EDT

PDB ID : 3AL4  
Title : Crystal structure of the swine-origin A (H1N1)-2009 influenza A virus hemagglutinin (HA) reveals similar antigenicity to that of the 1918 pandemic virus  
Authors : Zhang, W.; Qi, J.X.; Shi, Y.; Li, Q.; Yan, J.H.; Gao, G.F.  
Deposited on : 2010-07-22  
Resolution : 2.87 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 : 2.36  
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) : 2.36

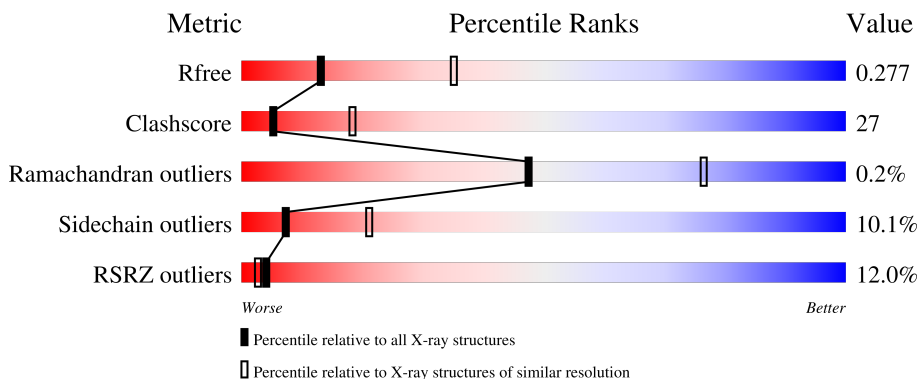
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.87 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	2691 (2.90-2.86)
Clashscore	141614	2947 (2.90-2.86)
Ramachandran outliers	138981	2868 (2.90-2.86)
Sidechain outliers	138945	2871 (2.90-2.86)
RSRZ outliers	127900	2629 (2.90-2.86)

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	333	 9% 54% 36% 6%
1	C	333	 8% 52% 41%
1	E	333	 14% 41% 47% 8%
1	G	333	 12% 54% 37% 6%
1	I	333	 8% 56% 35% 6%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	333	
2	B	181	
2	D	181	
2	F	181	
2	H	181	
2	J	181	
2	L	181	
3	M	2	
3	N	2	
3	P	2	
3	Q	2	
3	S	2	
4	O	3	
4	R	3	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NAG	M	2	-	-	-	X
3	NAG	N	2	-	-	-	X
3	NAG	Q	2	-	-	-	X
5	NAG	A	606	-	-	-	X
5	NAG	C	601	-	-	-	X
5	NAG	C	607	-	-	-	X
5	NAG	E	601	-	-	-	X
5	NAG	E	604	-	-	-	X
5	NAG	E	605	-	-	-	X
5	NAG	F	601	-	-	-	X
5	NAG	G	602	-	-	-	X
5	NAG	G	603	-	-	-	X
5	NAG	K	603	-	-	-	X

Continued on next page...

*Continued from previous page...*

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
5	NAG	K	604	-	-	-	X
5	NAG	L	601	-	-	X	X

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 23626 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	321	2505	1584	433	477	11	0	0	0
1	C	321	2509	1586	433	479	11	0	0	0
1	E	321	2509	1586	433	479	11	0	0	0
1	G	321	2505	1584	433	477	11	0	0	0
1	I	321	2509	1586	433	479	11	0	0	0
1	K	321	2511	1588	433	479	11	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	ALA	-	expression tag	UNP C3W5S1
A	2	ASP	-	expression tag	UNP C3W5S1
A	3	LEU	-	expression tag	UNP C3W5S1
A	4	GLY	-	expression tag	UNP C3W5S1
A	5	SER	-	expression tag	UNP C3W5S1
A	6	ARG	-	expression tag	UNP C3W5S1
C	1	ALA	-	expression tag	UNP C3W5S1
C	2	ASP	-	expression tag	UNP C3W5S1
C	3	LEU	-	expression tag	UNP C3W5S1
C	4	GLY	-	expression tag	UNP C3W5S1
C	5	SER	-	expression tag	UNP C3W5S1
C	6	ARG	-	expression tag	UNP C3W5S1
E	1	ALA	-	expression tag	UNP C3W5S1
E	2	ASP	-	expression tag	UNP C3W5S1
E	3	LEU	-	expression tag	UNP C3W5S1
E	4	GLY	-	expression tag	UNP C3W5S1
E	5	SER	-	expression tag	UNP C3W5S1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	6	ARG	-	expression tag	UNP C3W5S1
G	1	ALA	-	expression tag	UNP C3W5S1
G	2	ASP	-	expression tag	UNP C3W5S1
G	3	LEU	-	expression tag	UNP C3W5S1
G	4	GLY	-	expression tag	UNP C3W5S1
G	5	SER	-	expression tag	UNP C3W5S1
G	6	ARG	-	expression tag	UNP C3W5S1
I	1	ALA	-	expression tag	UNP C3W5S1
I	2	ASP	-	expression tag	UNP C3W5S1
I	3	LEU	-	expression tag	UNP C3W5S1
I	4	GLY	-	expression tag	UNP C3W5S1
I	5	SER	-	expression tag	UNP C3W5S1
I	6	ARG	-	expression tag	UNP C3W5S1
K	1	ALA	-	expression tag	UNP C3W5S1
K	2	ASP	-	expression tag	UNP C3W5S1
K	3	LEU	-	expression tag	UNP C3W5S1
K	4	GLY	-	expression tag	UNP C3W5S1
K	5	SER	-	expression tag	UNP C3W5S1
K	6	ARG	-	expression tag	UNP C3W5S1

- Molecule 2 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	162	1305	822	220	257	6	0	0	0
2	D	162	1300	818	219	257	6	0	0	0
2	F	161	1302	821	219	256	6	0	0	0
2	H	162	1305	822	220	257	6	0	0	0
2	J	162	1305	822	220	257	6	0	0	0
2	L	161	1302	821	219	256	6	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	177	ARG	-	expression tag	UNP C3W5S1
B	178	LEU	-	expression tag	UNP C3W5S1
B	179	VAL	-	expression tag	UNP C3W5S1

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	180	PRO	-	expression tag	UNP C3W5S1
B	181	ARG	-	expression tag	UNP C3W5S1
D	177	ARG	-	expression tag	UNP C3W5S1
D	178	LEU	-	expression tag	UNP C3W5S1
D	179	VAL	-	expression tag	UNP C3W5S1
D	180	PRO	-	expression tag	UNP C3W5S1
D	181	ARG	-	expression tag	UNP C3W5S1
F	177	ARG	-	expression tag	UNP C3W5S1
F	178	LEU	-	expression tag	UNP C3W5S1
F	179	VAL	-	expression tag	UNP C3W5S1
F	180	PRO	-	expression tag	UNP C3W5S1
F	181	ARG	-	expression tag	UNP C3W5S1
H	177	ARG	-	expression tag	UNP C3W5S1
H	178	LEU	-	expression tag	UNP C3W5S1
H	179	VAL	-	expression tag	UNP C3W5S1
H	180	PRO	-	expression tag	UNP C3W5S1
H	181	ARG	-	expression tag	UNP C3W5S1
J	177	ARG	-	expression tag	UNP C3W5S1
J	178	LEU	-	expression tag	UNP C3W5S1
J	179	VAL	-	expression tag	UNP C3W5S1
J	180	PRO	-	expression tag	UNP C3W5S1
J	181	ARG	-	expression tag	UNP C3W5S1
L	177	ARG	-	expression tag	UNP C3W5S1
L	178	LEU	-	expression tag	UNP C3W5S1
L	179	VAL	-	expression tag	UNP C3W5S1
L	180	PRO	-	expression tag	UNP C3W5S1
L	181	ARG	-	expression tag	UNP C3W5S1

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



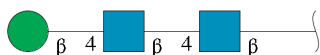
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	M	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	N	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	P	2	Total	C	N	O	0	0	0
			28	16	2	10			

Continued on next page...

Continued from previous page...

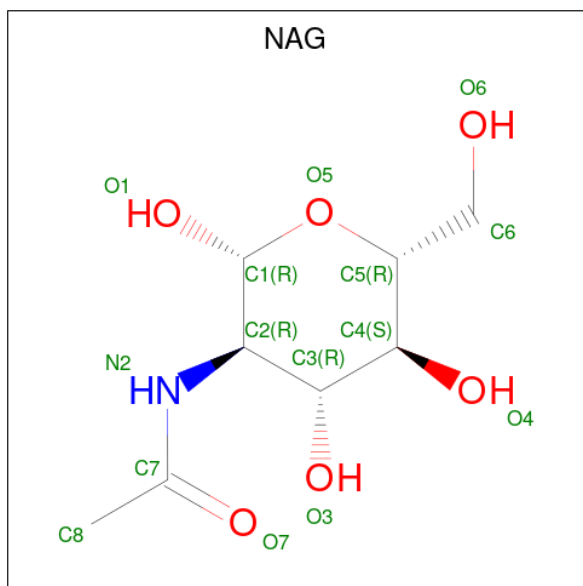
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	Q	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	S	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 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
			Total	C	N	O			
4	O	3	Total	C	N	O	0	0	0
			39	22	2	15			
4	R	3	Total	C	N	O	0	0	0
			39	22	2	15			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

Continued on next page...



*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total C N O 14 8 1 5	0	0
5	C	1	Total C N O 14 8 1 5	0	0
5	E	1	Total C N O 14 8 1 5	0	0
5	E	1	Total C N O 14 8 1 5	0	0
5	E	1	Total C N O 14 8 1 5	0	0
5	F	1	Total C N O 14 8 1 5	0	0
5	G	1	Total C N O 14 8 1 5	0	0
5	G	1	Total C N O 14 8 1 5	0	0
5	G	1	Total C N O 14 8 1 5	0	0
5	K	1	Total C N O 14 8 1 5	0	0
5	K	1	Total C N O 14 8 1 5	0	0
5	K	1	Total C N O 14 8 1 5	0	0
5	K	1	Total C N O 14 8 1 5	0	0
5	L	1	Total C N O 14 8 1 5	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	36	Total O 36 36	0	0
6	B	17	Total O 17 17	0	0
6	C	38	Total O 38 38	0	0
6	D	10	Total O 10 10	0	0
6	E	31	Total O 31 31	0	0

*Continued on next page...*

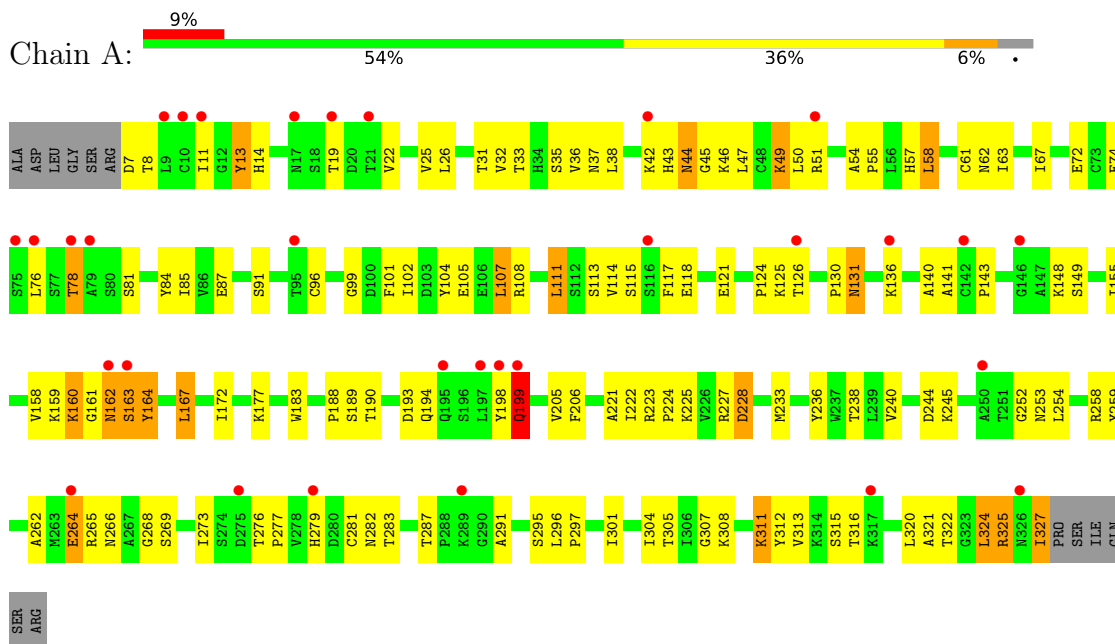
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	F	13	Total 13	O 13	0	0
6	G	45	Total 45	O 45	0	0
6	H	23	Total 23	O 23	0	0
6	I	37	Total 37	O 37	0	0
6	J	21	Total 21	O 21	0	0
6	K	35	Total 35	O 35	0	0
6	L	11	Total 11	O 11	0	0

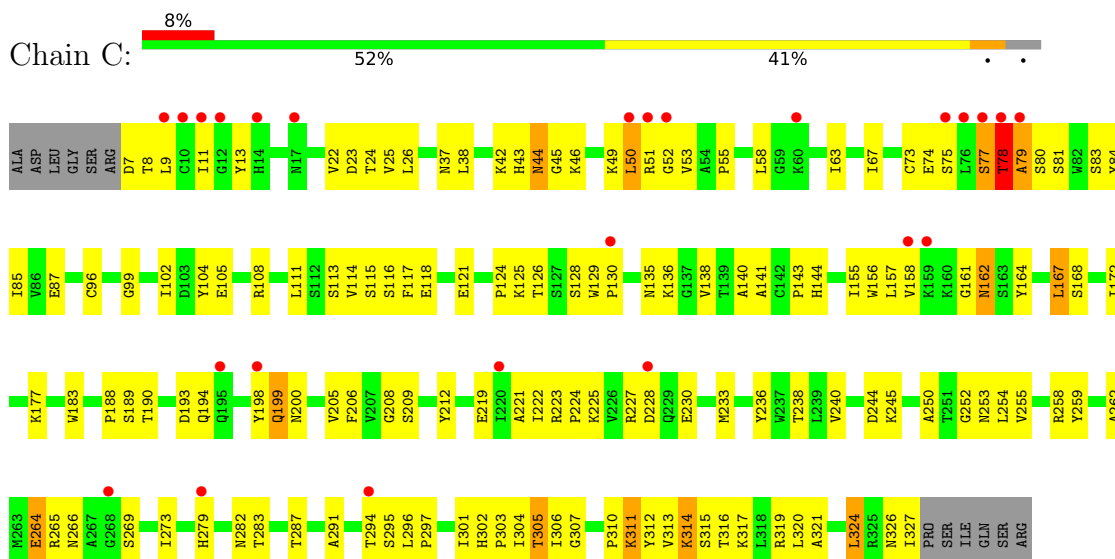
### 3 Residue-property plots [i](#)

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

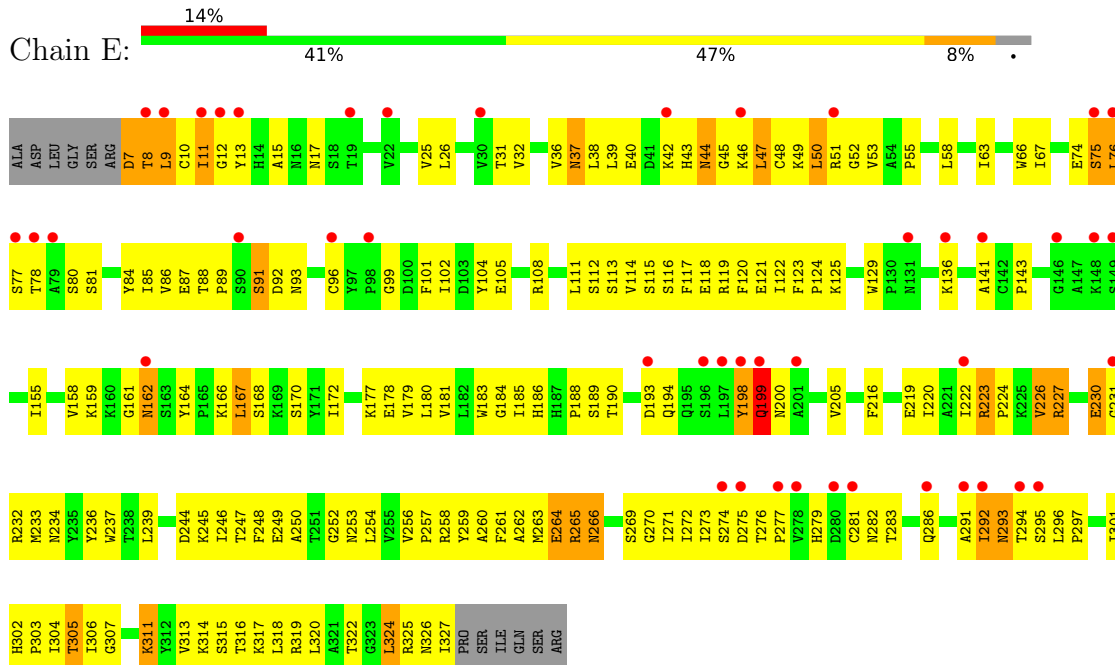
- Molecule 1: Hemagglutinin



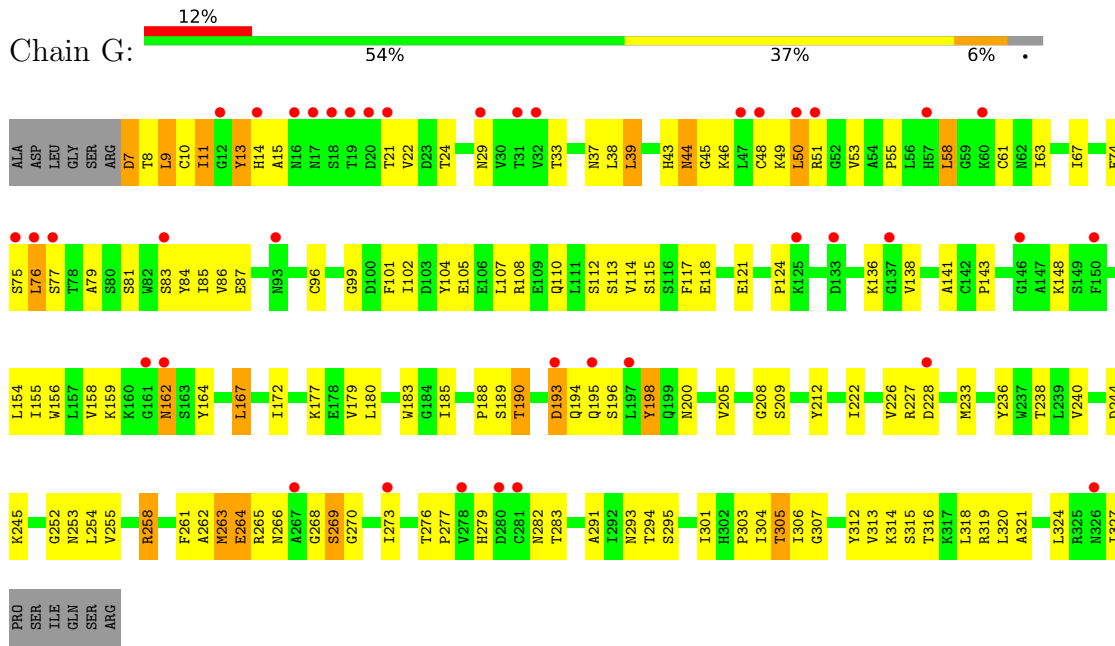
- Molecule 1: Hemagglutinin



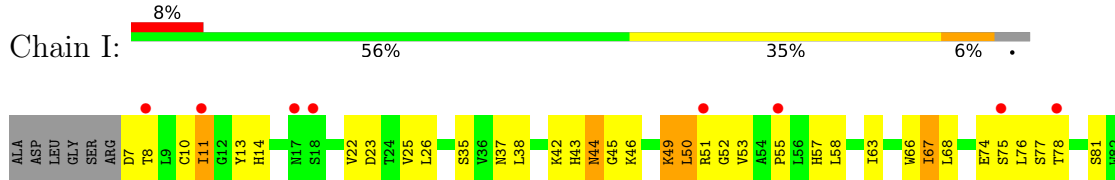
• Molecule 1: Hemagglutinin

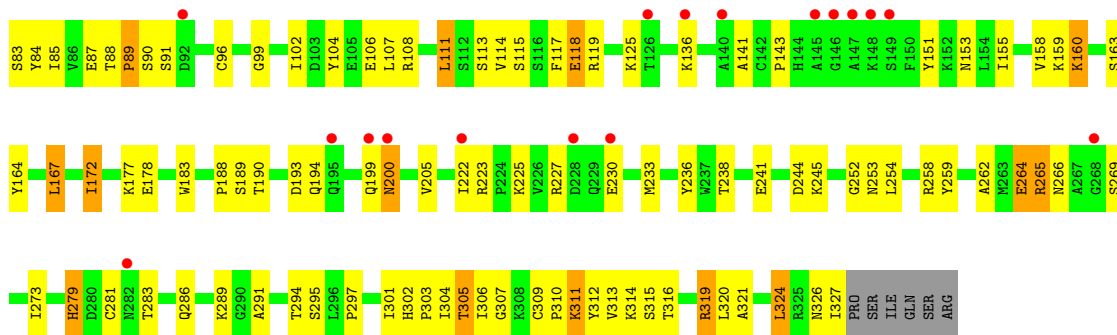


• Molecule 1: Hemagglutinin

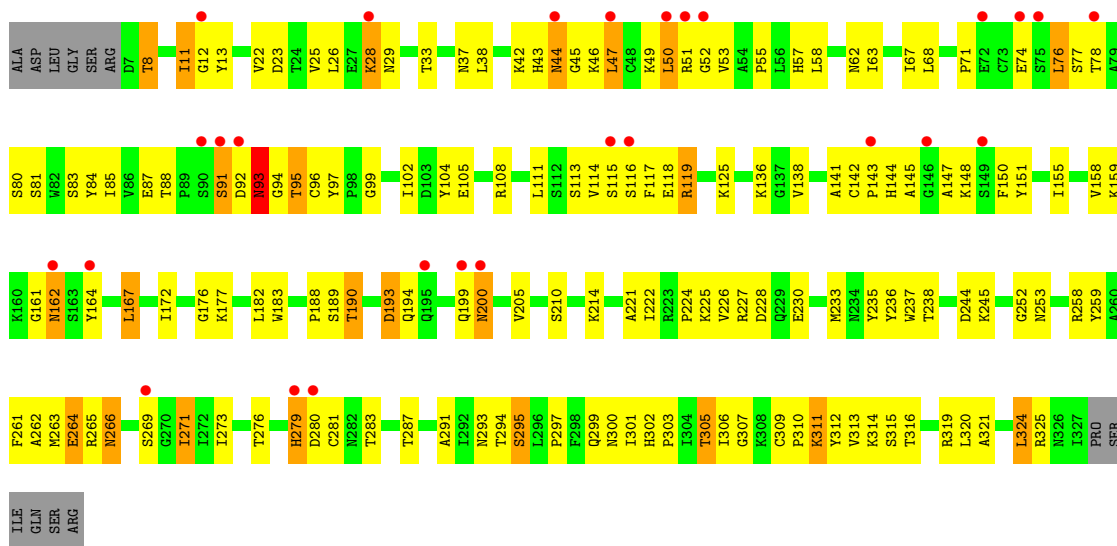


• Molecule 1: Hemagglutinin

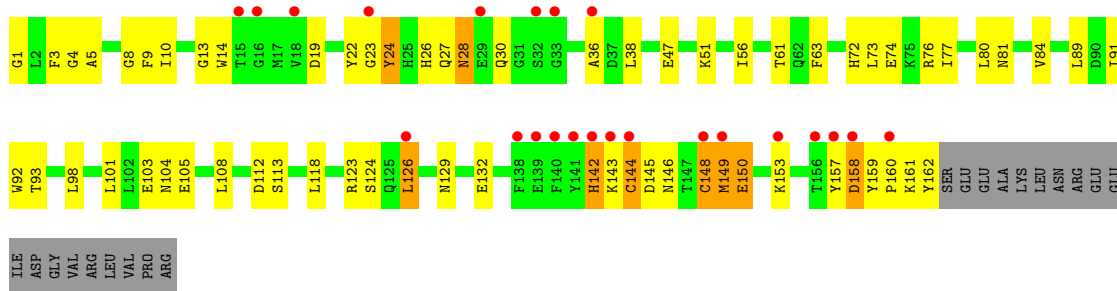




• Molecule 1: Hemagglutinin

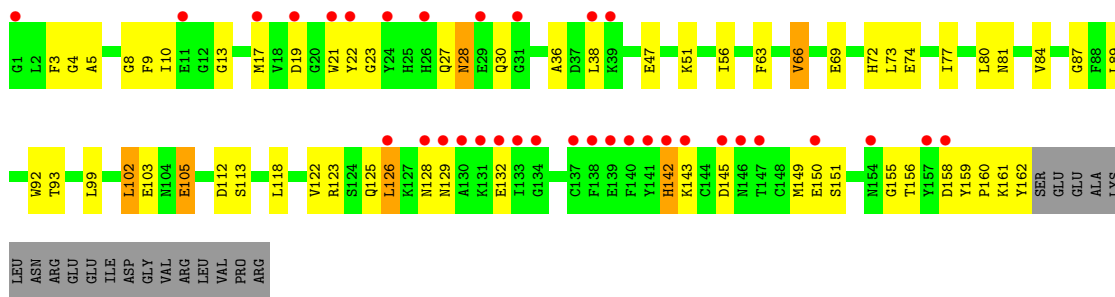


• Molecule 2: Hemagglutinin

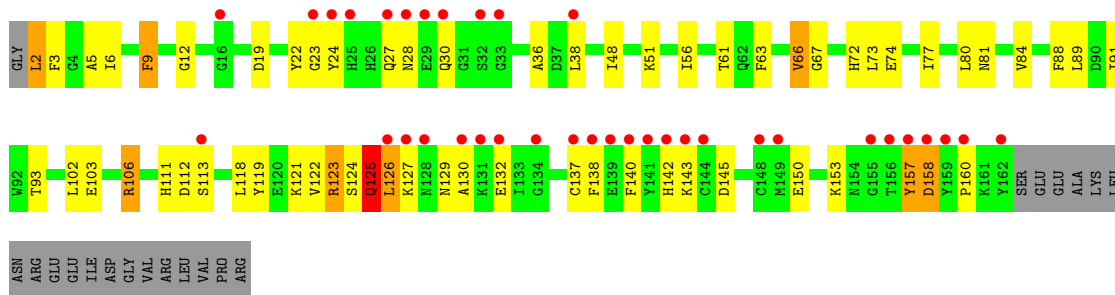


• Molecule 2: Hemagglutinin

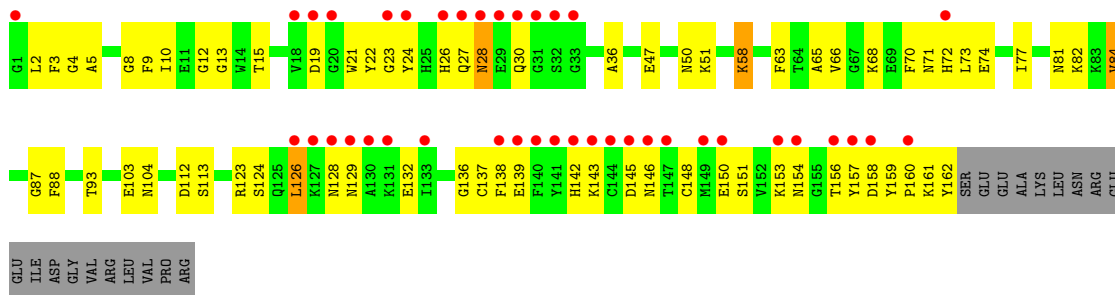




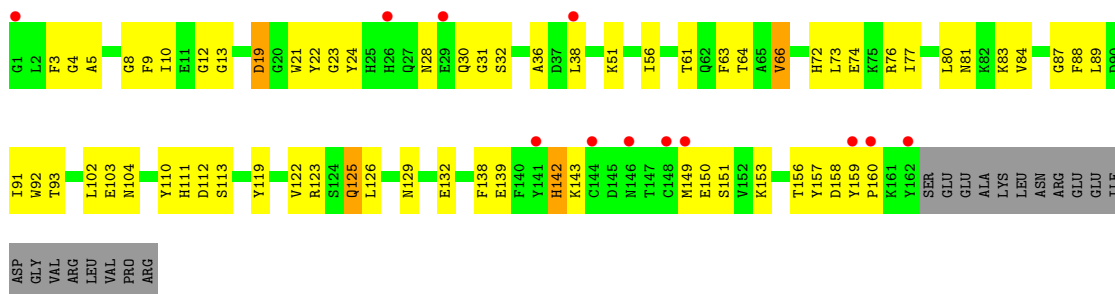
• Molecule 2: Hemagglutinin



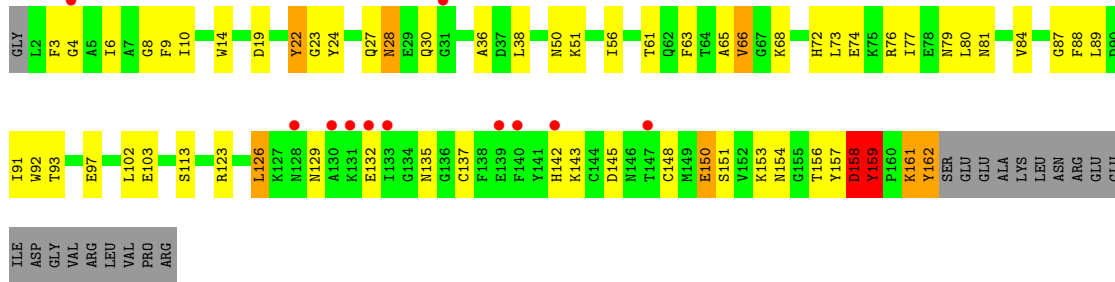
• Molecule 2: Hemagglutinin



• Molecule 2: Hemagglutinin



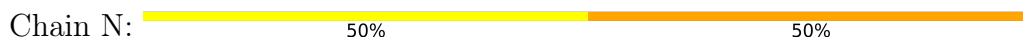
• Molecule 2: Hemagglutinin



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



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

Chain O:  33% 33% 33%

MAG1  
MAG2  
BMA3

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

Chain R:  33% 67%

MAG1  
MAG2  
BMA3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.02Å 115.19Å 114.98Å 62.31° 77.94° 81.05°	Depositor
Resolution (Å)	24.19 – 2.87 35.21 – 2.87	Depositor EDS
% Data completeness (in resolution range)	91.6 (24.19-2.87) 97.1 (35.21-2.87)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.53 (at 2.85Å)	Xtrriage
Refinement program	PHENIX (PHENIX.REFINE: 1.5_2)	Depositor
R, $R_{free}$	0.246 , 0.270 0.254 , 0.277	Depositor DCC
$R_{free}$ test set	3295 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	61.4	Xtrriage
Anisotropy	0.127	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 71.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	23626	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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: NAG, BMA

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	0/2568	0.56	5/3488 (0.1%)
1	C	0.26	0/2572	0.53	3/3493 (0.1%)
1	E	0.25	0/2572	0.62	7/3493 (0.2%)
1	G	0.26	0/2568	0.52	1/3488 (0.0%)
1	I	0.25	0/2572	0.51	3/3493 (0.1%)
1	K	0.28	0/2574	0.55	4/3497 (0.1%)
2	B	0.25	0/1333	0.43	0/1797
2	D	0.23	0/1328	0.38	0/1791
2	F	0.24	0/1330	0.45	1/1794 (0.1%)
2	H	0.25	0/1333	0.43	0/1797
2	J	0.25	0/1333	0.40	0/1797
2	L	0.28	0/1330	0.55	3/1794 (0.2%)
All	All	0.26	0/23413	0.52	27/31722 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	2
1	E	0	1
2	F	0	1
2	L	0	1
All	All	0	5

There are no bond length outliers.

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	226	VAL	O-C-N	-14.34	99.75	122.70
1	A	199	GLN	N-CA-CB	-11.55	89.81	110.60
2	L	159	TYR	N-CA-CB	-11.34	90.19	110.60
1	E	76	LEU	N-CA-CB	-10.33	89.75	110.40
1	E	226	VAL	CA-C-N	9.25	137.55	117.20

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	73	CYS	Peptide
1	C	77	SER	Peptide
1	E	226	VAL	Mainchain
2	F	125	GLN	Peptide
2	L	159	TYR	Peptide

## 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	2505	0	2442	156	0
1	C	2509	0	2447	133	3
1	E	2509	0	2449	213	0
1	G	2505	0	2443	141	0
1	I	2509	0	2450	136	0
1	K	2511	0	2454	164	3
2	B	1305	0	1228	71	0
2	D	1300	0	1216	54	0
2	F	1302	0	1224	70	0
2	H	1305	0	1228	92	0
2	J	1305	0	1228	77	0
2	L	1302	0	1225	73	0
3	M	28	0	25	9	0
3	N	28	0	25	4	0
3	P	28	0	25	2	0
3	Q	28	0	25	3	0
3	S	28	0	25	5	0
4	O	39	0	34	3	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	R	39	0	34	3	0
5	A	28	0	26	0	0
5	C	28	0	26	3	0
5	E	42	0	39	2	0
5	F	14	0	13	0	0
5	G	42	0	39	3	0
5	K	56	0	52	9	0
5	L	14	0	13	10	0
6	A	36	0	0	19	0
6	B	17	0	0	11	0
6	C	38	0	0	15	0
6	D	10	0	0	2	0
6	E	31	0	0	10	0
6	F	13	0	0	4	0
6	G	45	0	0	25	0
6	H	23	0	0	13	0
6	I	37	0	0	21	0
6	J	21	0	0	7	0
6	K	35	0	0	15	0
6	L	11	0	0	2	0
All	All	23626	0	22435	1215	3

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 27.

The worst 5 of 1215 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:17:ASN:HD21	5:E:601:NAG:C1	1.30	1.42
1:E:7:ASP:HA	2:F:27:GLN:O	1.37	1.21
1:E:66:TRP:HE1	1:E:77:SER:HB2	1.08	1.14
1:K:280:ASP:OD1	3:M:2:NAG:O7	1.65	1.14
1:C:282:ASN:OD1	5:C:607:NAG:O7	1.67	1.13

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:74:GLU:OE1	1:K:144:HIS:NE2[1_455]	1.89	0.31
1:C:74:GLU:OE1	1:K:144:HIS:CD2[1_455]	2.04	0.16

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:74:GLU:OE2	1:K:144:HIS:ND1[1_455]	2.13	0.07

## 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	319/333 (96%)	286 (90%)	32 (10%)	1 (0%)	41	70
1	C	319/333 (96%)	296 (93%)	22 (7%)	1 (0%)	41	70
1	E	319/333 (96%)	291 (91%)	24 (8%)	4 (1%)	12	35
1	G	319/333 (96%)	299 (94%)	19 (6%)	1 (0%)	41	70
1	I	319/333 (96%)	294 (92%)	25 (8%)	0	100	100
1	K	319/333 (96%)	288 (90%)	31 (10%)	0	100	100
2	B	160/181 (88%)	143 (89%)	17 (11%)	0	100	100
2	D	160/181 (88%)	146 (91%)	14 (9%)	0	100	100
2	F	159/181 (88%)	147 (92%)	12 (8%)	0	100	100
2	H	160/181 (88%)	145 (91%)	15 (9%)	0	100	100
2	J	160/181 (88%)	148 (92%)	12 (8%)	0	100	100
2	L	159/181 (88%)	145 (91%)	14 (9%)	0	100	100
All	All	2872/3084 (93%)	2628 (92%)	237 (8%)	7 (0%)	47	76

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	92	ASP
1	E	199	GLN
1	A	199	GLN
1	E	75	SER
1	G	268	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	281/292 (96%)	249 (89%)	32 (11%)	5	16
1	C	282/292 (97%)	255 (90%)	27 (10%)	8	23
1	E	282/292 (97%)	247 (88%)	35 (12%)	4	12
1	G	281/292 (96%)	252 (90%)	29 (10%)	7	20
1	I	282/292 (97%)	251 (89%)	31 (11%)	6	17
1	K	282/292 (97%)	249 (88%)	33 (12%)	5	14
2	B	139/156 (89%)	126 (91%)	13 (9%)	8	24
2	D	138/156 (88%)	128 (93%)	10 (7%)	14	37
2	F	139/156 (89%)	124 (89%)	15 (11%)	6	18
2	H	139/156 (89%)	132 (95%)	7 (5%)	24	54
2	J	139/156 (89%)	127 (91%)	12 (9%)	10	29
2	L	139/156 (89%)	127 (91%)	12 (9%)	10	29
All	All	2523/2688 (94%)	2267 (90%)	256 (10%)	7	21

5 of 256 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	205	VAL
1	K	287	THR
1	E	167	LEU
1	E	91	SER
1	K	324	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 94 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	30	GLN
2	J	27	GLN
2	H	129	ASN
1	I	144	HIS

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
2	J	125	GLN

### 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	$\# Z  > 2$
3	NAG	M	1	1,3	14,14,15	0.55	0	17,19,21	1.07	0
3	NAG	M	2	3	14,14,15	0.57	0	17,19,21	1.55	3 (17%)
3	NAG	N	1	1,3	14,14,15	0.57	0	17,19,21	0.65	0
3	NAG	N	2	3	14,14,15	0.46	0	17,19,21	1.05	1 (5%)
4	NAG	O	1	4,1	14,14,15	0.61	0	17,19,21	1.15	1 (5%)
4	NAG	O	2	4	14,14,15	0.55	0	17,19,21	1.12	1 (5%)
4	BMA	O	3	4	11,11,12	0.26	0	15,15,17	0.64	0
3	NAG	P	1	1,3	14,14,15	0.55	0	17,19,21	0.92	1 (5%)
3	NAG	P	2	3	14,14,15	0.44	0	17,19,21	2.01	4 (23%)
3	NAG	Q	1	1,3	14,14,15	0.56	0	17,19,21	0.66	0
3	NAG	Q	2	3	14,14,15	0.58	0	17,19,21	0.81	0
4	NAG	R	1	4,1	14,14,15	0.53	0	17,19,21	1.71	3 (17%)
4	NAG	R	2	4	14,14,15	0.40	0	17,19,21	1.88	4 (23%)
4	BMA	R	3	4	11,11,12	0.27	0	15,15,17	0.66	0
3	NAG	S	1	1,3	14,14,15	0.54	0	17,19,21	1.81	4 (23%)
3	NAG	S	2	3	14,14,15	0.61	0	17,19,21	1.13	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
3	NAG	M	1	1,3	-	4/6/23/26	0/1/1/1
3	NAG	M	2	3	-	4/6/23/26	0/1/1/1
3	NAG	N	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	N	2	3	-	2/6/23/26	0/1/1/1
4	NAG	O	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	O	2	4	-	0/6/23/26	0/1/1/1
4	BMA	O	3	4	-	2/2/19/22	0/1/1/1
3	NAG	P	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	P	2	3	-	4/6/23/26	0/1/1/1
3	NAG	Q	1	1,3	-	3/6/23/26	0/1/1/1
3	NAG	Q	2	3	-	2/6/23/26	0/1/1/1
4	NAG	R	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	R	2	4	-	0/6/23/26	0/1/1/1
4	BMA	R	3	4	-	2/2/19/22	0/1/1/1
3	NAG	S	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	S	2	3	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	P	2	NAG	C1-O5-C5	6.16	120.54	112.19
4	R	2	NAG	C1-O5-C5	5.09	119.08	112.19
3	S	1	NAG	O4-C4-C3	4.96	121.81	110.35
4	R	1	NAG	C2-N2-C7	-4.92	115.90	122.90
3	M	2	NAG	C2-N2-C7	-4.29	116.80	122.90

There are no chirality outliers.

5 of 29 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	M	2	NAG	C4-C5-C6-O6
3	M	1	NAG	C4-C5-C6-O6
3	M	2	NAG	O5-C5-C6-O6
3	M	1	NAG	O5-C5-C6-O6

*Continued on next page...*



*Continued from previous page...*

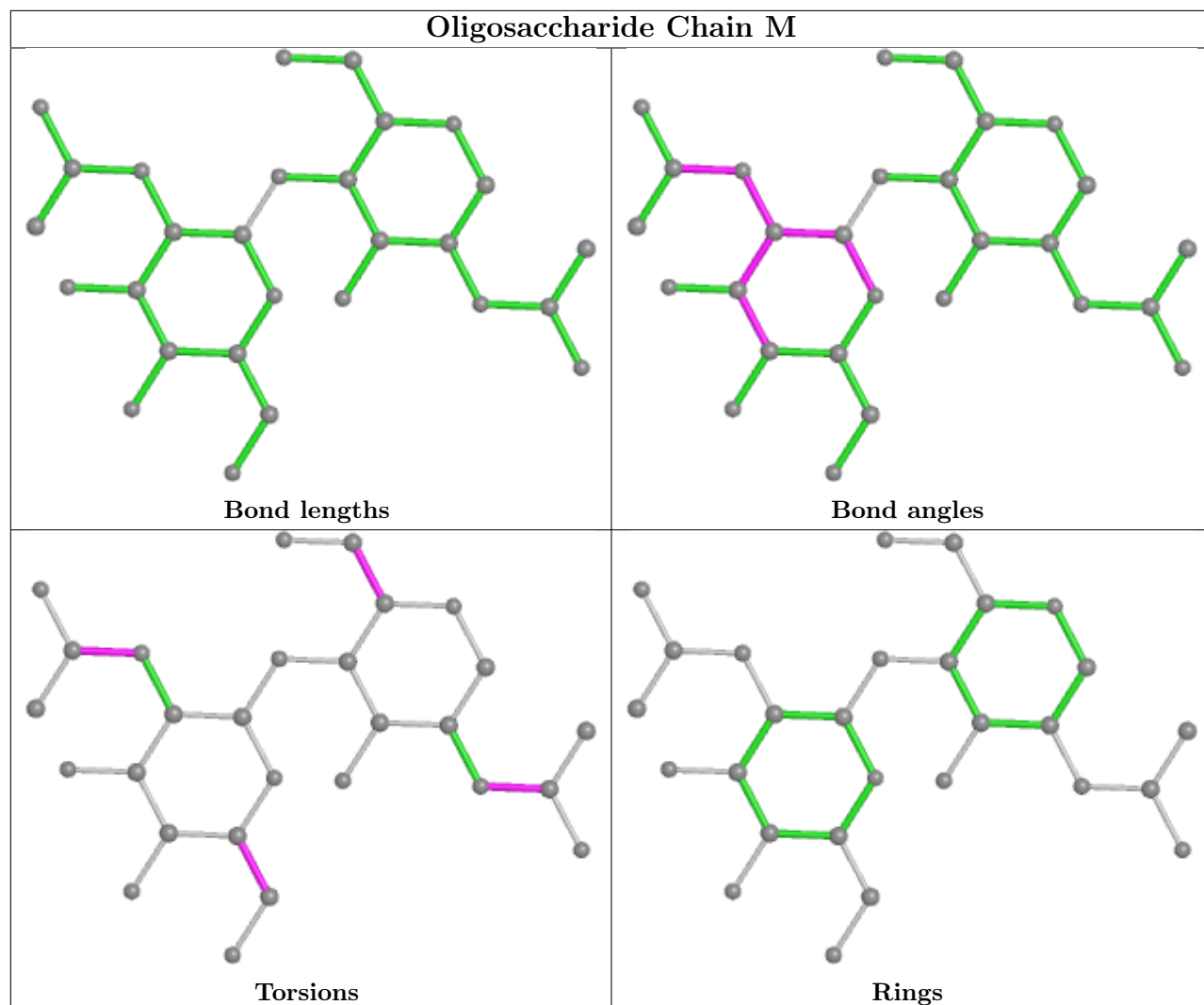
Mol	Chain	Res	Type	Atoms
4	O	1	NAG	O5-C5-C6-O6

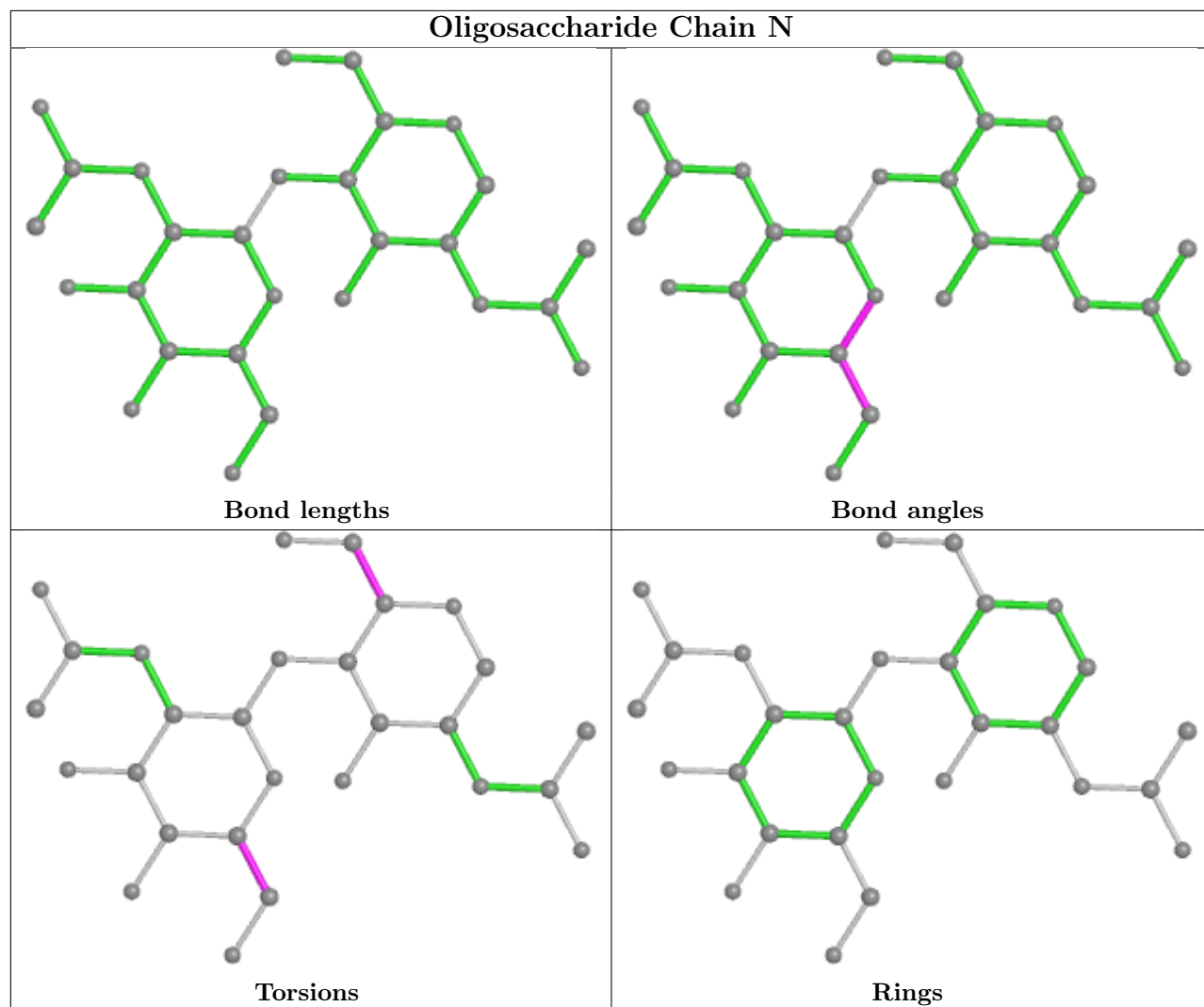
There are no ring outliers.

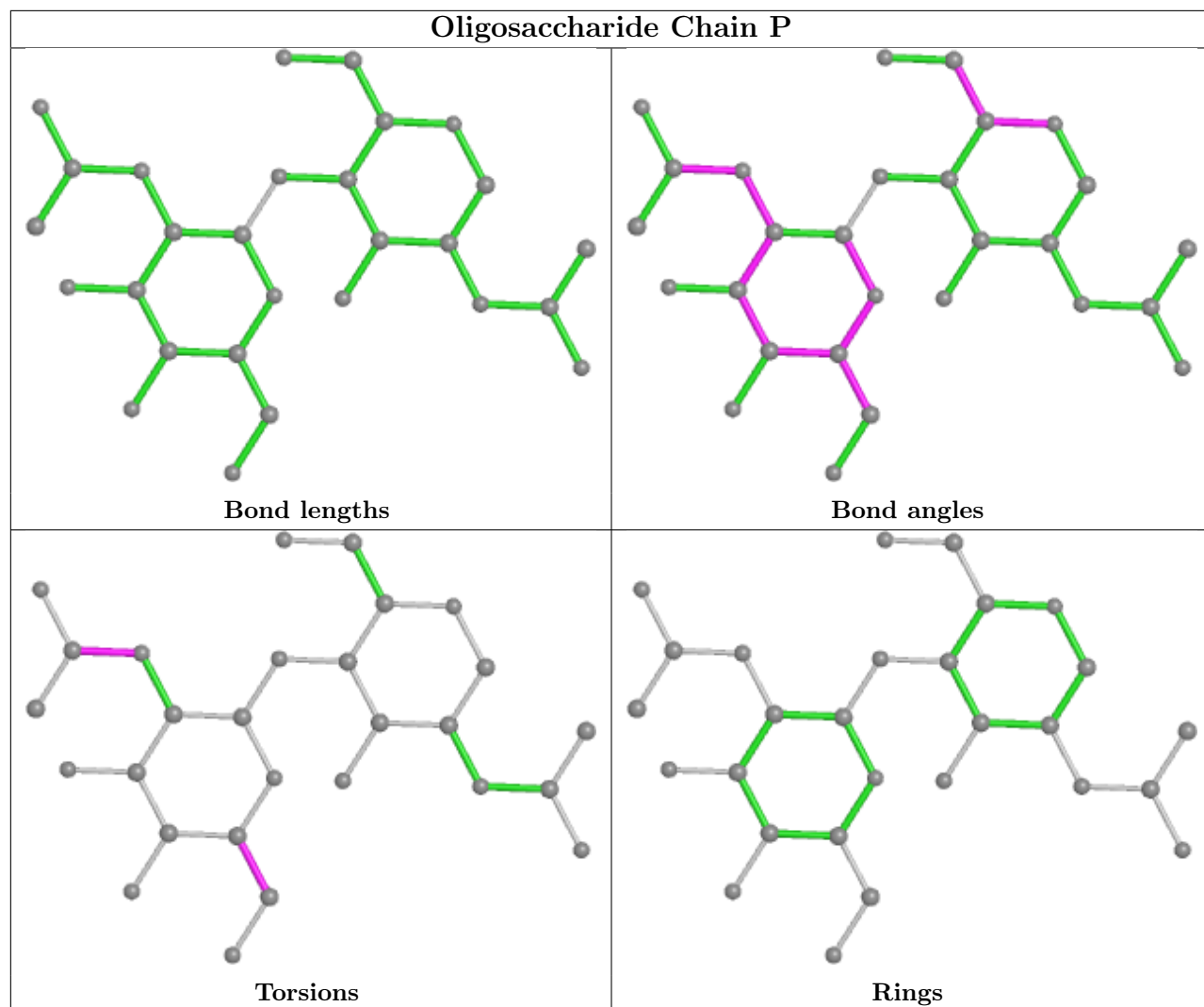
12 monomers are involved in 29 short contacts:

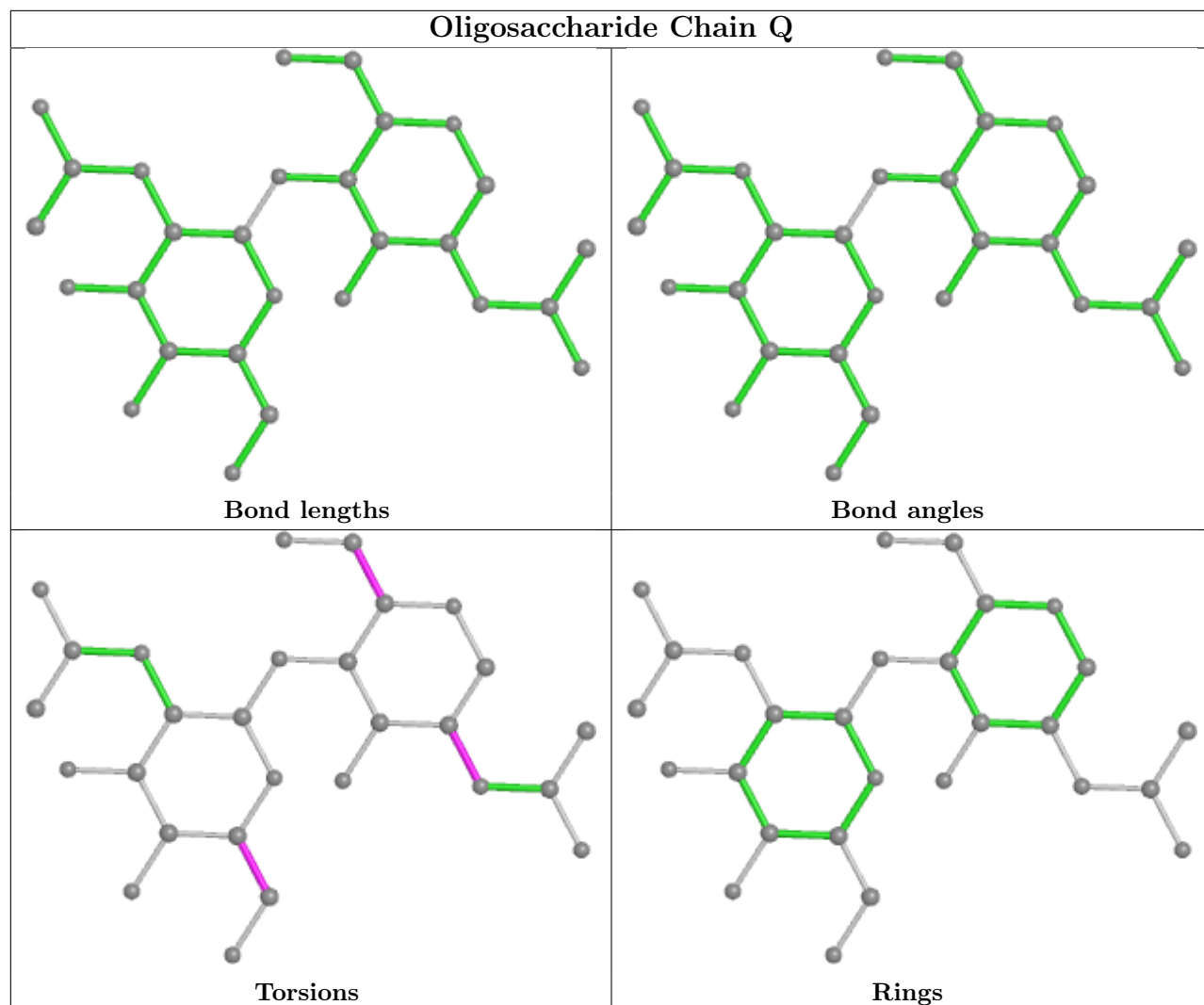
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	N	1	NAG	4	0
4	R	2	NAG	2	0
3	M	1	NAG	6	0
4	O	1	NAG	3	0
3	S	2	NAG	3	0
3	P	1	NAG	2	0
3	M	2	NAG	3	0
3	Q	1	NAG	3	0
4	R	1	NAG	1	0
4	R	3	BMA	2	0
3	S	1	NAG	3	0
3	N	2	NAG	1	0

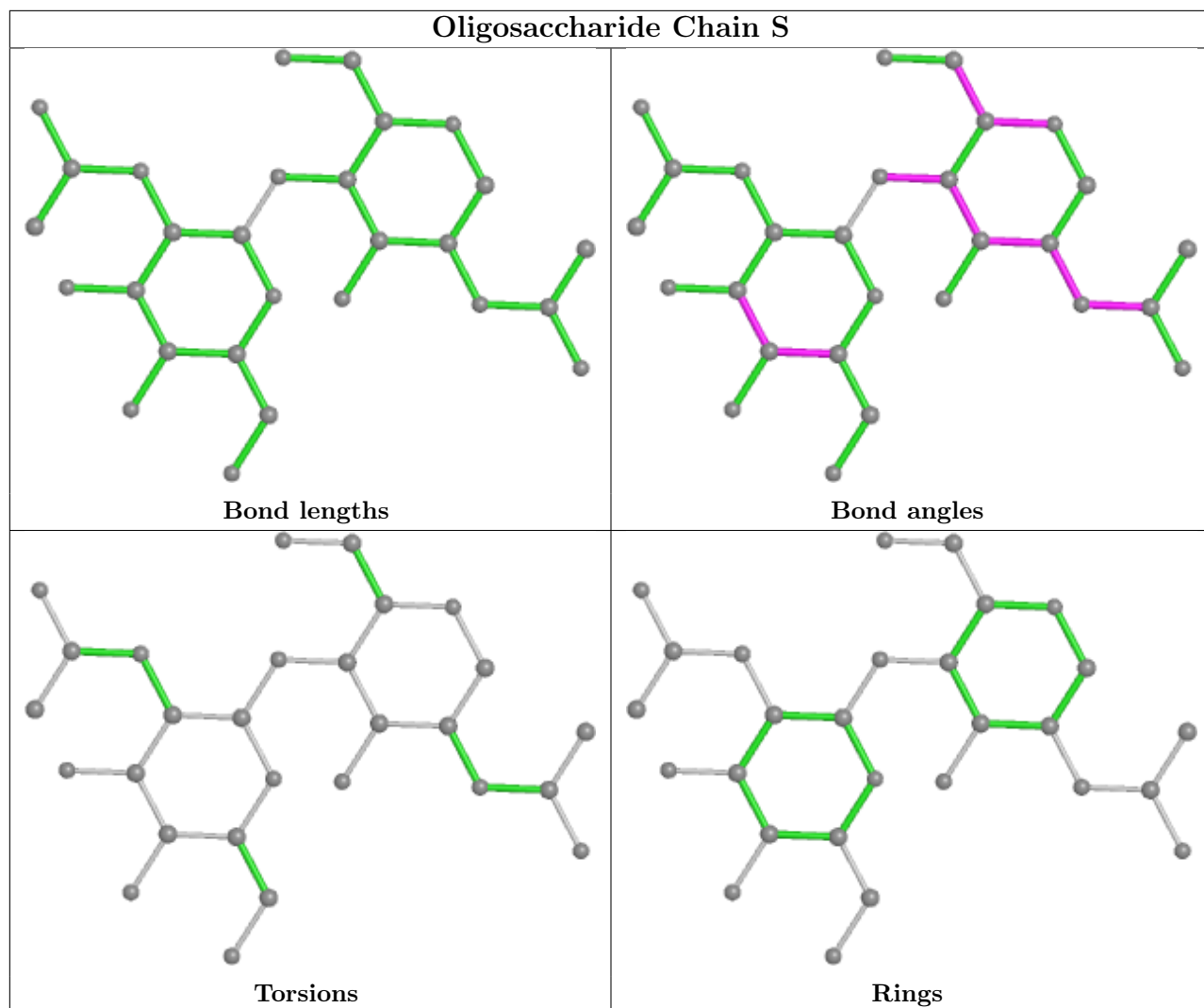
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

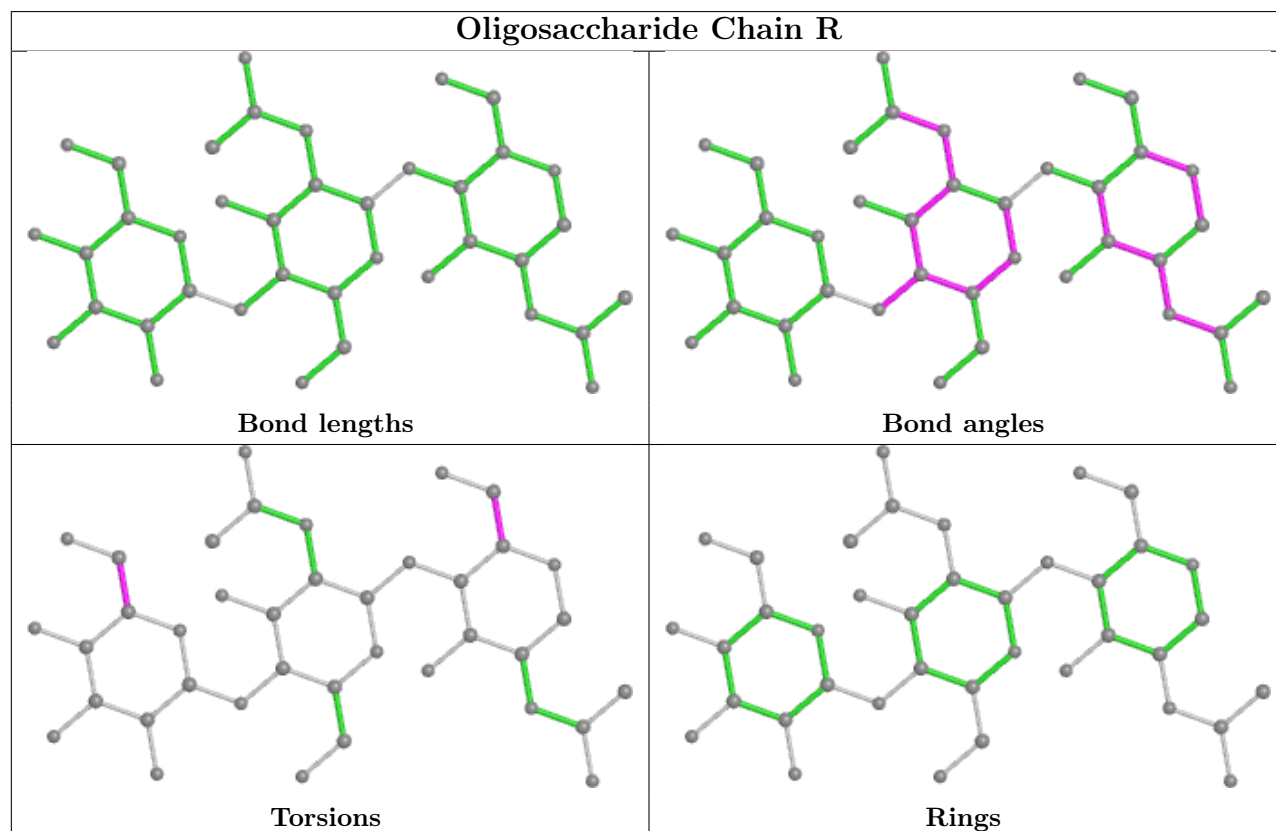
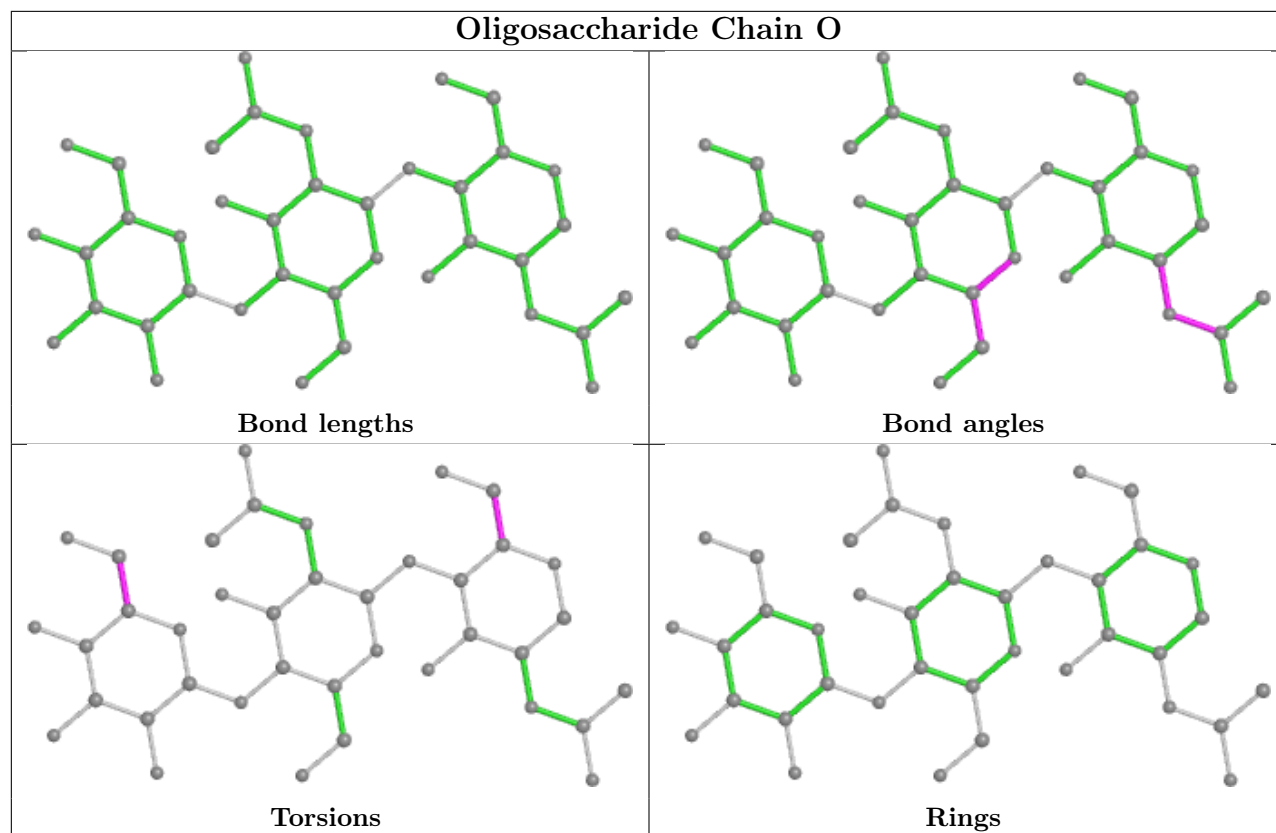












## 5.6 Ligand geometry

16 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	A	606	1	14,14,15	0.55	0	17,19,21	0.92	0
5	NAG	C	601	1	14,14,15	0.57	0	17,19,21	0.66	0
5	NAG	L	601	2	14,14,15	0.58	0	17,19,21	0.66	0
5	NAG	C	607	1	14,14,15	0.56	0	17,19,21	0.66	0
5	NAG	G	601	1	14,14,15	0.54	0	17,19,21	0.82	1 (5%)
5	NAG	E	601	1	14,14,15	0.57	0	17,19,21	0.65	0
5	NAG	F	601	2	14,14,15	0.56	0	17,19,21	0.66	0
5	NAG	K	603	1	14,14,15	0.57	0	17,19,21	0.66	0
5	NAG	G	603	1	14,14,15	0.57	0	17,19,21	0.66	0
5	NAG	A	601	1	14,14,15	0.56	0	17,19,21	0.66	0
5	NAG	E	605	1	14,14,15	0.56	0	17,19,21	1.33	5 (29%)
5	NAG	K	601	1	14,14,15	0.46	0	17,19,21	2.46	4 (23%)
5	NAG	G	602	1	14,14,15	0.56	0	17,19,21	0.66	0
5	NAG	K	602	1	14,14,15	0.56	0	17,19,21	0.67	0
5	NAG	K	604	1	14,14,15	0.57	0	17,19,21	0.67	0
5	NAG	E	604	1	14,14,15	0.54	0	17,19,21	0.72	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
5	NAG	A	606	1	-	2/6/23/26	0/1/1/1
5	NAG	C	601	1	-	0/6/23/26	0/1/1/1
5	NAG	L	601	2	-	2/6/23/26	0/1/1/1
5	NAG	C	607	1	-	2/6/23/26	0/1/1/1
5	NAG	G	601	1	-	2/6/23/26	0/1/1/1
5	NAG	E	601	1	-	2/6/23/26	0/1/1/1
5	NAG	F	601	2	-	2/6/23/26	0/1/1/1

*Continued on next page...*



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	K	603	1	-	2/6/23/26	0/1/1/1
5	NAG	G	603	1	-	3/6/23/26	0/1/1/1
5	NAG	A	601	1	-	2/6/23/26	0/1/1/1
5	NAG	E	605	1	-	2/6/23/26	0/1/1/1
5	NAG	K	601	1	-	4/6/23/26	0/1/1/1
5	NAG	G	602	1	-	2/6/23/26	0/1/1/1
5	NAG	K	602	1	-	2/6/23/26	0/1/1/1
5	NAG	K	604	1	-	1/6/23/26	0/1/1/1
5	NAG	E	604	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	K	601	NAG	C1-O5-C5	8.73	124.03	112.19
5	K	601	NAG	O5-C1-C2	3.11	116.20	111.29
5	E	605	NAG	O5-C1-C2	-2.68	107.05	111.29
5	E	605	NAG	C6-C5-C4	-2.43	107.32	113.00
5	G	601	NAG	O5-C5-C6	2.25	110.72	107.20

There are no chirality outliers.

5 of 30 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	606	NAG	C8-C7-N2-C2
5	A	606	NAG	O7-C7-N2-C2
5	K	602	NAG	C4-C5-C6-O6
5	G	601	NAG	O5-C5-C6-O6
5	K	601	NAG	C4-C5-C6-O6

There are no ring outliers.

8 monomers are involved in 27 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	L	601	NAG	10	0
5	C	607	NAG	3	0
5	E	601	NAG	2	0
5	K	603	NAG	3	0
5	G	603	NAG	1	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	602	NAG	2	0
5	K	602	NAG	4	0
5	K	604	NAG	2	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	321/333 (96%)	0.56	31 (9%) 7 5	39, 70, 122, 226	0
1	C	321/333 (96%)	0.41	25 (7%) 13 10	36, 67, 116, 196	0
1	E	321/333 (96%)	0.87	45 (14%) 2 2	41, 74, 130, 215	0
1	G	321/333 (96%)	0.71	39 (12%) 4 2	46, 76, 128, 232	0
1	I	321/333 (96%)	0.54	25 (7%) 13 10	37, 70, 120, 228	0
1	K	321/333 (96%)	0.43	27 (8%) 11 8	38, 70, 121, 195	0
2	B	162/181 (89%)	0.86	23 (14%) 2 2	32, 77, 167, 209	0
2	D	162/181 (89%)	1.02	34 (20%) 1 0	34, 72, 186, 246	0
2	F	161/181 (88%)	1.02	36 (22%) 0 0	34, 74, 178, 241	0
2	H	162/181 (89%)	1.19	40 (24%) 0 0	33, 80, 174, 289	0
2	J	162/181 (89%)	0.61	12 (7%) 14 11	32, 69, 165, 200	0
2	L	161/181 (88%)	0.56	11 (6%) 17 13	36, 73, 169, 199	0
All	All	2896/3084 (93%)	0.68	348 (12%) 4 3	32, 72, 157, 289	0

The worst 5 of 348 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	159	TYR	13.4
2	F	128	ASN	9.8
2	D	139	GLU	9.4
2	H	31	GLY	8.7
2	D	140	PHE	8.4

### 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

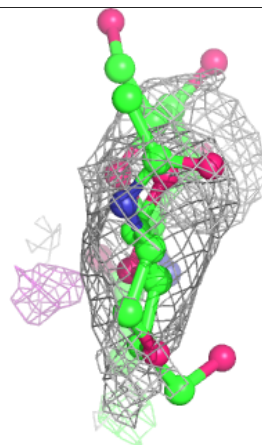
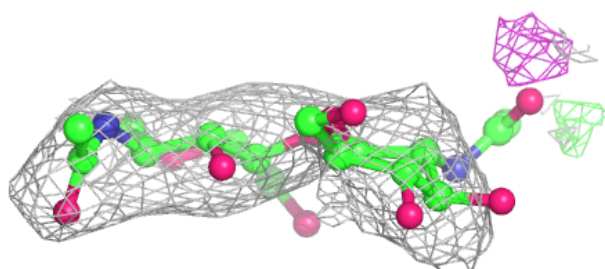
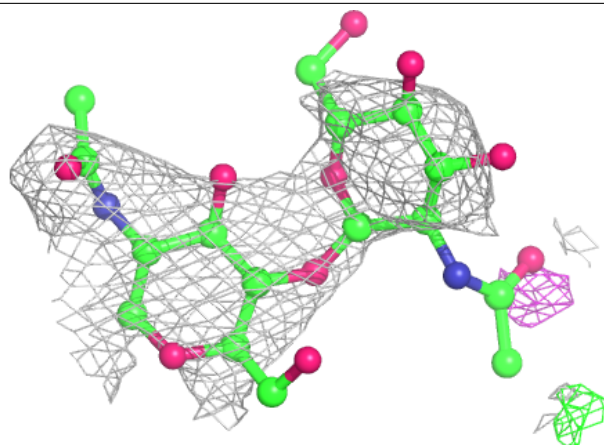
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	BMA	R	3	11/12	0.46	0.30	108,135,142,143	0
3	NAG	N	2	14/15	0.61	0.40	124,134,147,148	0
3	NAG	M	2	14/15	0.67	0.55	120,148,155,174	0
3	NAG	Q	2	14/15	0.73	0.52	119,139,166,193	0
4	BMA	O	3	11/12	0.77	0.27	89,112,124,151	0
3	NAG	P	2	14/15	0.77	0.33	91,106,136,144	0
3	NAG	S	2	14/15	0.78	0.30	128,134,141,147	0
3	NAG	N	1	14/15	0.79	0.34	68,108,117,126	0
3	NAG	Q	1	14/15	0.79	0.27	108,129,134,135	0
3	NAG	S	1	14/15	0.80	0.27	94,114,139,150	0
4	NAG	R	1	14/15	0.81	0.21	72,82,95,97	0
4	NAG	O	2	14/15	0.82	0.28	55,84,111,120	0
3	NAG	M	1	14/15	0.87	0.40	82,113,129,130	0
3	NAG	P	1	14/15	0.89	0.19	70,86,98,99	0
4	NAG	O	1	14/15	0.89	0.14	61,76,95,102	0
4	NAG	R	2	14/15	0.91	0.33	92,117,138,145	0

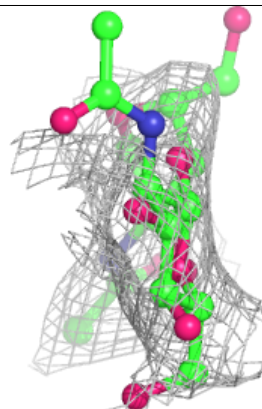
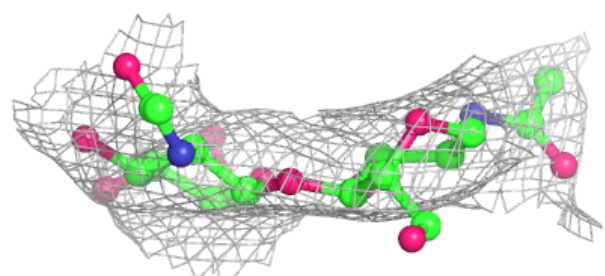
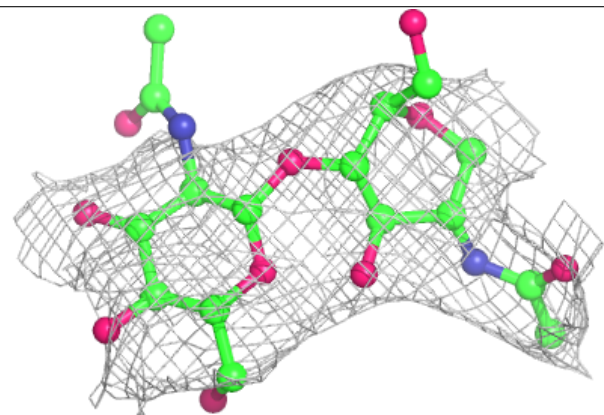
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain M:**

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

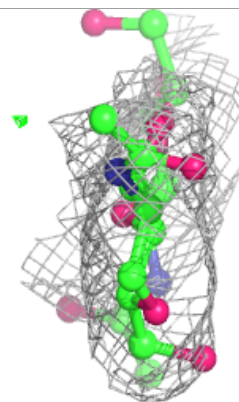
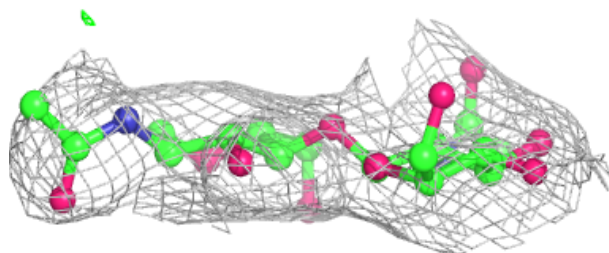
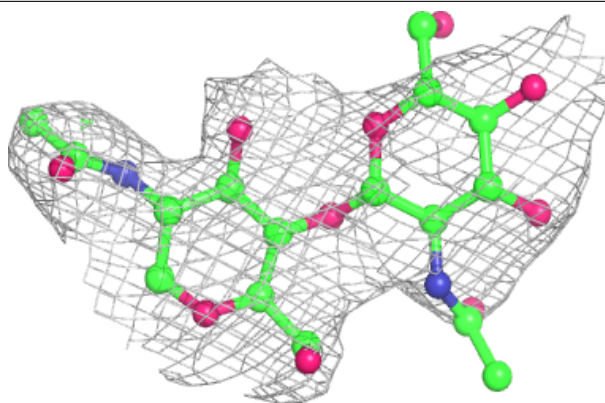
**Electron density around Chain N:**

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

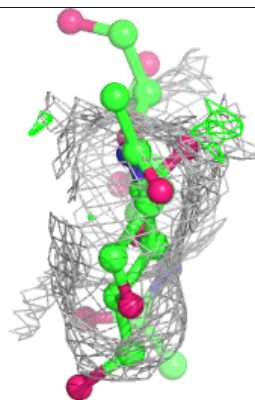
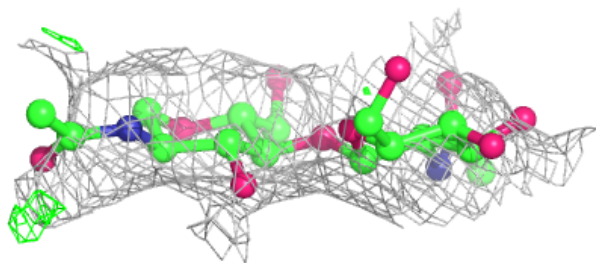
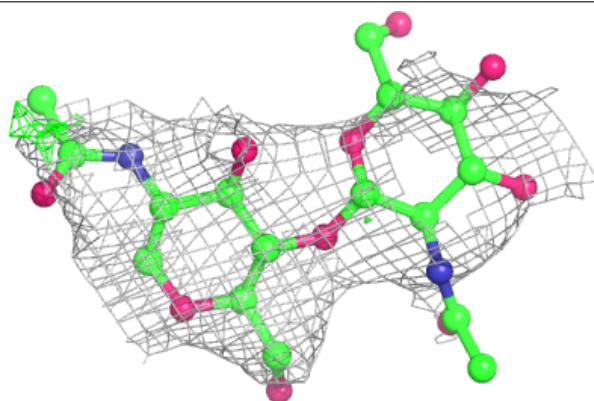


**Electron density around Chain P:**

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

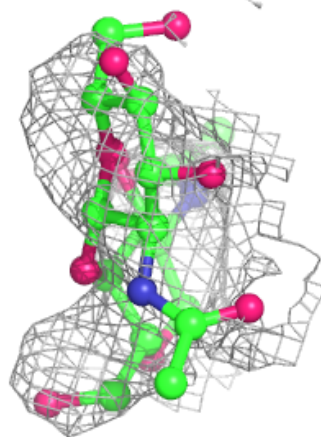
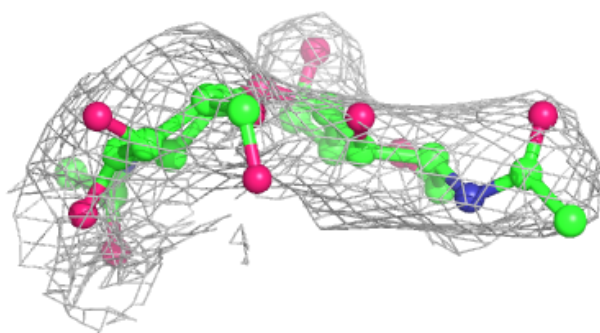
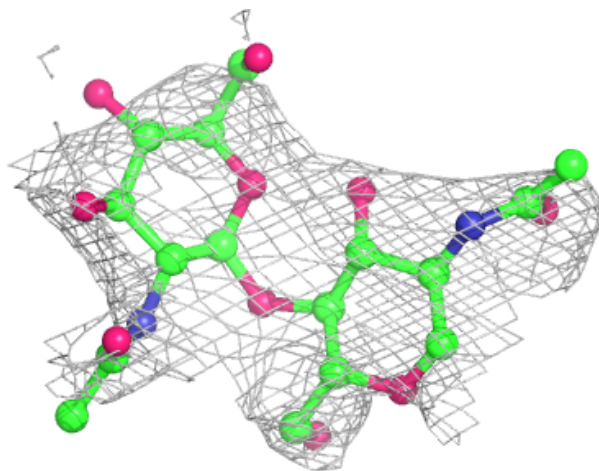
**Electron density around Chain Q:**

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



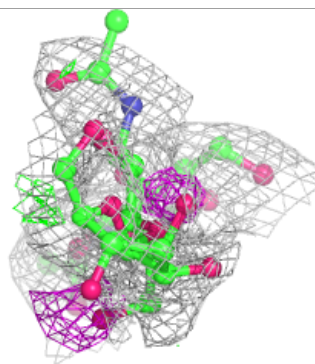
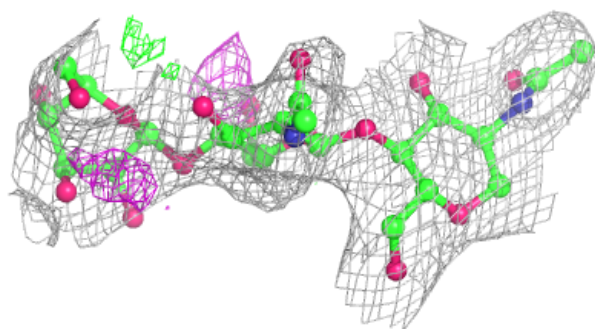
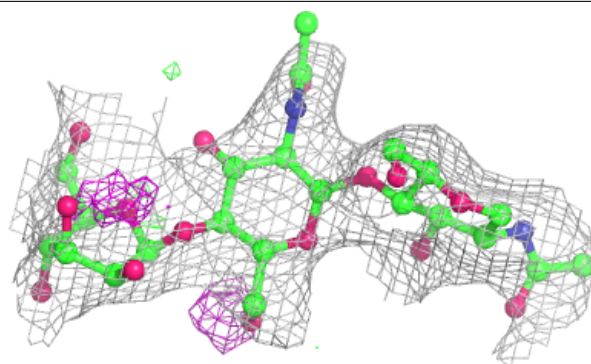
**Electron density around Chain S:**

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

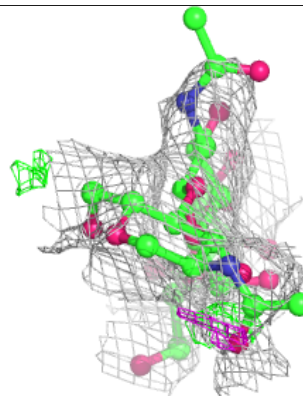
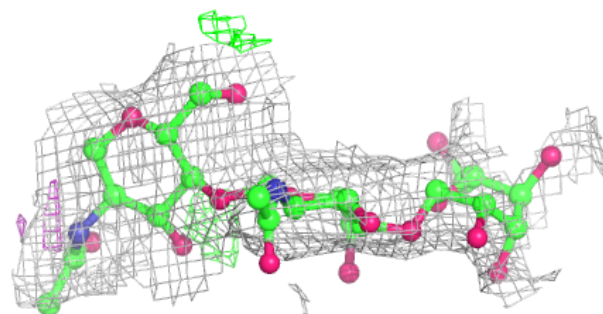
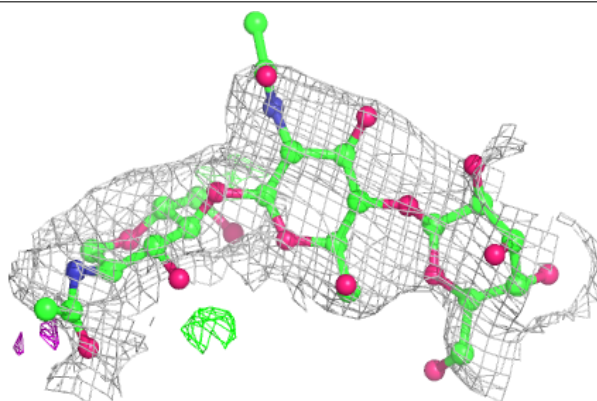


**Electron density around Chain O:**

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

**Electron density around Chain R:**

$2mF_o-DF_c$  (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	L	601	14/15	0.31	0.54	132,161,178,187	0
5	NAG	E	601	14/15	0.48	0.53	117,156,185,194	0
5	NAG	A	606	14/15	0.51	0.51	103,126,141,143	0
5	NAG	E	605	14/15	0.54	0.59	134,171,193,199	0
5	NAG	K	601	14/15	0.56	0.36	88,113,140,142	0
5	NAG	F	601	14/15	0.58	0.44	154,173,188,194	0
5	NAG	K	602	14/15	0.63	0.34	68,115,138,139	0
5	NAG	A	601	14/15	0.64	0.34	53,78,94,104	0
5	NAG	C	601	14/15	0.66	0.50	157,180,192,198	0
5	NAG	C	607	14/15	0.66	0.42	95,116,142,146	0
5	NAG	G	602	14/15	0.67	0.49	171,186,198,202	0
5	NAG	G	601	14/15	0.69	0.33	141,158,176,181	0
5	NAG	K	603	14/15	0.72	0.51	84,126,147,153	0
5	NAG	K	604	14/15	0.72	0.50	108,133,146,153	0
5	NAG	E	604	14/15	0.72	0.53	26,28,30,30	0
5	NAG	G	603	14/15	0.75	0.41	139,163,185,199	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.