



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2020 – 12:42 PM BST

PDB ID : 5KUY  
Title : Influenza hemagglutinin H3 A/Hong Kong/1/1968 in complex with designed inhibitor protein HSB.2A  
Authors : Bernard, S.M.; Wilson, I.A.  
Deposited on : 2016-07-13  
Resolution : 2.60 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

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

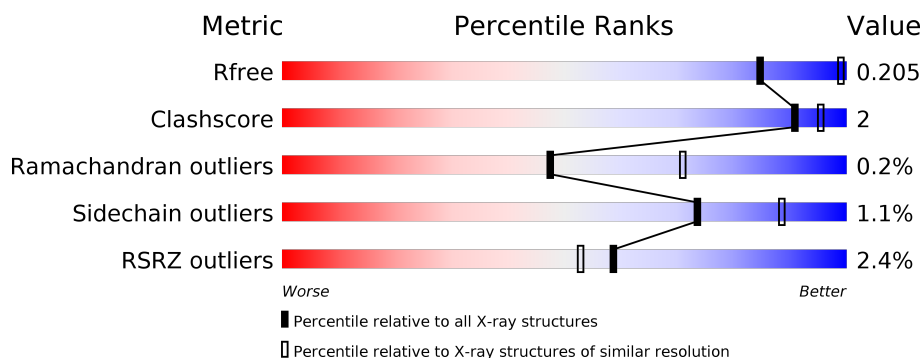
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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.






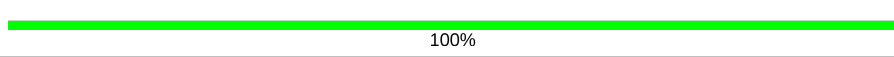

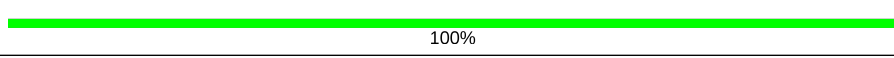
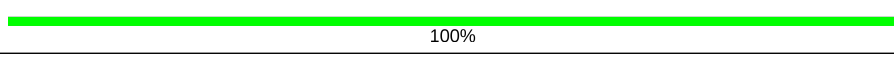
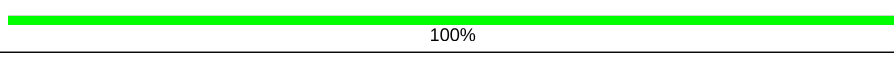
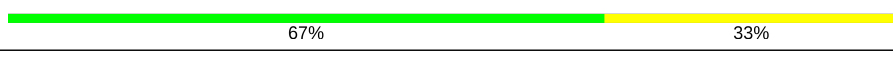
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	<div> <div>94%</div> <div>5%</div> </div>
1	C	323	<div> <div>92%</div> <div>6%</div> </div>
1	E	323	<div> <div>89%</div> <div>9%</div> </div>
2	B	177	<div> <div>98%</div> </div>
2	D	177	<div> <div>%</div> <div>94%</div> </div>
2	F	177	<div> <div>94%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	97	 % 78% 8% 13%
3	H	97	 2% 81% 6% 12%
3	I	97	 35% 53% 45%
4	J	2	 100%
4	K	2	 100%
4	L	2	 100%
4	M	2	 100%
4	O	2	 100%
5	N	3	 67% 33%

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 13967 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Hemagglutinin HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	319	Total	C	N	O	S	0	0	0
			2456	1538	431	474	13			
1	C	316	Total	C	N	O	S	0	0	0
			2434	1526	428	467	13			
1	E	317	Total	C	N	O	S	0	0	0
			2443	1531	429	470	13			

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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	176	Total	C	N	O	S	0	0	0
			1419	882	248	283	6			
2	D	172	Total	C	N	O	S	0	0	0
			1391	863	243	279	6			
2	F	171	Total	C	N	O	S	0	0	0
			1382	858	241	277	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	452	GLY	ARG	engineered mutation	UNP Q91MA7
B	506	SER	-	expression tag	UNP Q91MA7
D	452	GLY	ARG	engineered mutation	UNP Q91MA7
D	506	SER	-	expression tag	UNP Q91MA7
F	452	GLY	ARG	engineered mutation	UNP Q91MA7
F	506	SER	-	expression tag	UNP Q91MA7

- Molecule 3 is a protein called Designed influenza inhibitor HSB.2A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	84	Total	C	N	O	S	0	0	0
			660	421	109	128	2			
3	H	85	Total	C	N	O	S	0	0	0
			666	424	110	130	2			
3	I	53	Total	C	N	O	S	0	0	0
			383	248	60	74	1			

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



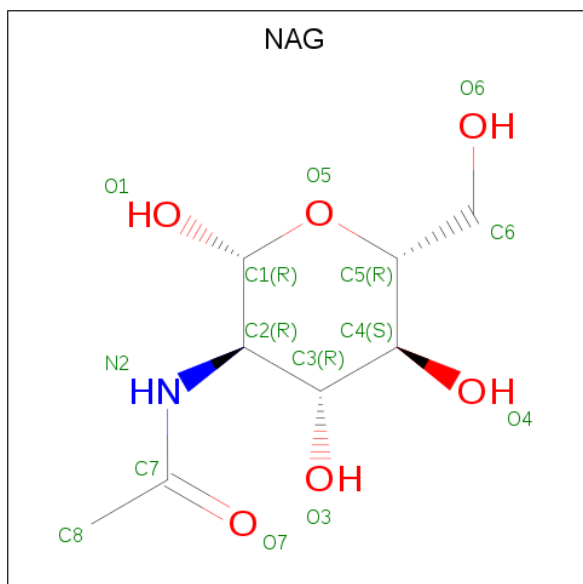
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	J	2	Total	C	N	O		0	0	0
			28	16	2	10				
4	K	2	Total	C	N	O		0	0	0
			28	16	2	10				
4	L	2	Total	C	N	O		0	0	0
			28	16	2	10				
4	M	2	Total	C	N	O		0	0	0
			28	16	2	10				
4	O	2	Total	C	N	O		0	0	0
			28	16	2	10				

- 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	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	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	B	1	Total	C	N	O	0	0
			14	8	1	5		
6	D	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	100	Total 100	O 100	0	0
7	B	71	Total 71	O 71	0	0
7	C	60	Total 60	O 60	0	0
7	D	59	Total 59	O 59	0	0
7	E	69	Total 69	O 69	0	0
7	F	78	Total 78	O 78	0	0
7	G	17	Total 17	O 17	0	0
7	H	16	Total 16	O 16	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 HA1

Chain A: 




- Molecule 1: Hemagglutinin HA1

Chain C: 



- Molecule 1: Hemagglutinin HA1

Chain E: 

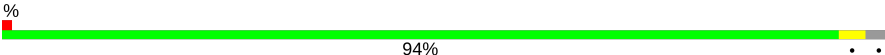


- Molecule 2: Hemagglutinin HA2

Chain B: 



- Molecule 2: Hemagglutinin HA2

Chain D: 






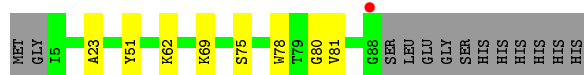
- Molecule 2: Hemagglutinin HA2

Chain F:  94%




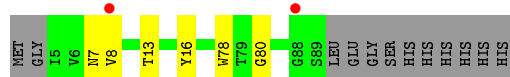
- Molecule 3: Designed influenza inhibitor HSB.2A

Chain G:  78%



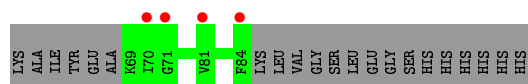
- Molecule 3: Designed influenza inhibitor HSB.2A

Chain H:  81%



- Molecule 3: Designed influenza inhibitor HSB.2A

Chain I:  53%



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

Chain J:  100%



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

Chain K:  100%



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

Chain L:  100%

HA01  
HA02

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

Chain M:  100%

HA01  
HA02

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

Chain O:  100%

HA01  
HA02

- 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:  67% 33%

HA01  
HA02  
E/A3

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	154.26Å 219.32Å 104.52Å 90.00° 117.78° 90.00°	Depositor
Resolution (Å)	48.13 – 2.60 48.14 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.13-2.60) 96.0 (48.14-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.01 (at 2.61Å)	Xtriage
Refinement program	PHENIX (1.10 _2155: ???)	Depositor
R, $R_{free}$	0.173 , 0.204 0.173 , 0.205	Depositor DCC
$R_{free}$ test set	4712 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.1	Xtriage
Anisotropy	0.596	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 34.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13967	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% 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	0.26	0/2513	0.47	0/3426
1	C	0.24	0/2491	0.46	0/3395
1	E	0.25	0/2500	0.46	0/3407
2	B	0.25	0/1443	0.42	0/1939
2	D	0.25	0/1415	0.43	0/1902
2	F	0.25	0/1406	0.44	0/1890
3	G	0.26	0/673	0.46	0/915
3	H	0.26	0/679	0.45	0/923
3	I	0.23	0/389	0.41	0/531
All	All	0.25	0/13509	0.45	0/18328

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	2456	0	2399	8	0
1	C	2434	0	2385	11	0
1	E	2443	0	2391	13	0
2	B	1419	0	1339	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1391	0	1303	4	0
2	F	1382	0	1295	4	0
3	G	660	0	638	3	0
3	H	666	0	643	3	0
3	I	383	0	322	1	0
4	J	28	0	25	0	0
4	K	28	0	25	0	0
4	L	28	0	25	0	0
4	M	28	0	25	0	0
4	O	28	0	25	0	0
5	N	39	0	34	1	0
6	A	28	0	26	0	0
6	B	14	0	13	0	0
6	D	14	0	13	0	0
6	E	28	0	26	0	0
7	A	100	0	0	0	0
7	B	71	0	0	0	0
7	C	60	0	0	0	0
7	D	59	0	0	0	0
7	E	69	0	0	0	0
7	F	78	0	0	1	0
7	G	17	0	0	0	0
7	H	16	0	0	0	0
All	All	13967	0	12952	44	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 2.

All (44) 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:119:GLU:OE1	1:E:261:ARG:NH2	2.26	0.68
1:E:77:ASP:OD2	1:E:141:ARG:NH1	2.29	0.65
1:E:283:THR:HG22	1:E:301:THR:HG22	1.79	0.63
1:C:57:ARG:NH1	1:C:83:THR:O	2.34	0.60
1:A:283:THR:HG22	1:A:301:THR:HG22	1.86	0.57
2:B:383:ARG:HH12	1:C:27:LYS:HD3	1.72	0.55
1:E:102:VAL:HG22	1:E:232:ILE:HB	1.91	0.53
2:D:456:ARG:NH1	2:F:460:GLU:OE1	2.42	0.53
1:A:102:VAL:HG22	1:A:232:ILE:HB	1.90	0.52
3:H:13:THR:HG23	3:H:16:TYR:H	1.74	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:195:TYR:O	1:C:197:GLN:N	2.42	0.52
1:C:102:VAL:HG22	1:C:232:ILE:HB	1.90	0.52
2:F:454:GLN:NE2	7:F:701:HOH:O	2.39	0.52
1:C:141:ARG:NH2	1:C:147:PHE:O	2.38	0.51
1:E:97:CYS:SG	1:E:98:TYR:N	2.83	0.50
1:A:195:TYR:O	1:A:197:GLN:N	2.44	0.49
1:C:27:LYS:HG2	1:C:32:ASP:O	2.12	0.49
1:A:27:LYS:HD3	2:F:383:ARG:HH12	1.78	0.49
1:C:236:ILE:HG12	1:C:260:MET:HE1	1.95	0.47
1:E:167:THR:HB	5:N:1:NAG:H62	1.97	0.46
3:I:27:TYR:CE1	3:I:31:GLU:HG3	2.52	0.45
1:C:97:CYS:SG	1:C:98:TYR:N	2.87	0.45
1:A:97:CYS:SG	1:A:98:TYR:N	2.88	0.45
2:D:497:ASN:OD1	2:D:501:GLN:NE2	2.49	0.45
2:B:383:ARG:NH1	1:C:27:LYS:HD3	2.32	0.45
1:E:77:ASP:CG	1:E:141:ARG:HH12	2.19	0.44
1:A:27:LYS:HG2	1:A:32:ASP:O	2.18	0.44
3:G:23:ALA:HA	3:G:81:VAL:HG21	1.99	0.44
1:A:27:LYS:HD3	2:F:383:ARG:NH1	2.32	0.43
1:E:87:PHE:HB3	1:E:267:ILE:HG13	2.00	0.43
1:E:50:LYS:HB3	1:E:275:ASP:HB2	1.99	0.43
1:E:182:VAL:HG21	1:E:213:ILE:HB	2.00	0.43
2:B:453:ARG:HD3	2:D:463:GLY:HA2	2.01	0.43
3:G:51:TYR:OH	3:G:69:LYS:NZ	2.51	0.43
1:C:117:THR:HG21	1:C:261:ARG:HH11	1.83	0.42
1:E:195:TYR:O	1:E:197:GLN:N	2.49	0.42
1:A:89:GLU:HG3	1:A:267:ILE:HD11	2.01	0.42
3:G:78:TRP:CZ2	3:G:80:GLY:HA3	2.53	0.42
1:C:49:GLY:HA2	1:C:285:ASN:O	2.20	0.42
1:E:156:LYS:NZ	1:E:193:SER:O	2.40	0.42
3:H:13:THR:HG22	3:H:16:TYR:HB2	2.02	0.42
2:D:460:GLU:HG3	2:D:499:ARG:HD2	2.01	0.41
1:E:222:TRP:CE2	1:E:225:GLY:HA2	2.55	0.41
3:H:78:TRP:CZ2	3:H:80:GLY:HA3	2.56	0.41

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	317/323 (98%)	309 (98%)	7 (2%)	1 (0%)	41	64
1	C	314/323 (97%)	306 (98%)	7 (2%)	1 (0%)	41	64
1	E	315/323 (98%)	306 (97%)	8 (2%)	1 (0%)	41	64
2	B	174/177 (98%)	166 (95%)	8 (5%)	0	100	100
2	D	170/177 (96%)	163 (96%)	7 (4%)	0	100	100
2	F	169/177 (96%)	161 (95%)	8 (5%)	0	100	100
3	G	82/97 (84%)	80 (98%)	2 (2%)	0	100	100
3	H	83/97 (86%)	82 (99%)	1 (1%)	0	100	100
3	I	45/97 (46%)	43 (96%)	2 (4%)	0	100	100
All	All	1669/1791 (93%)	1616 (97%)	50 (3%)	3 (0%)	47	71

All (3) Ramachandran outliers are listed below:

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

### 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	279/283 (99%)	276 (99%)	3 (1%)	73	88

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	277/283 (98%)	275 (99%)	2 (1%)	84	94
1	E	278/283 (98%)	272 (98%)	6 (2%)	52	76
2	B	149/150 (99%)	149 (100%)	0	100	100
2	D	146/150 (97%)	146 (100%)	0	100	100
2	F	145/150 (97%)	144 (99%)	1 (1%)	84	94
3	G	68/79 (86%)	66 (97%)	2 (3%)	42	68
3	H	69/79 (87%)	67 (97%)	2 (3%)	42	68
3	I	32/79 (40%)	32 (100%)	0	100	100
All	All	1443/1536 (94%)	1427 (99%)	16 (1%)	73	88

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	243	LEU
1	A	260	MET
1	C	18	HIS
1	C	260	MET
1	E	18	HIS
1	E	32	ASP
1	E	54	ASN
1	E	186	SER
1	E	260	MET
1	E	325	GLU
2	F	356	GLN
3	G	62	LYS
3	G	75	SER
3	H	7	ASN
3	H	8	VAL

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

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.



## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

13 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	J	1	1,4	14,14,15	0.37	0	17,19,21	0.39	0
4	NAG	J	2	4	14,14,15	0.36	0	17,19,21	0.55	0
4	NAG	K	1	1,4	14,14,15	0.36	0	17,19,21	0.43	0
4	NAG	K	2	4	14,14,15	0.23	0	17,19,21	0.56	0
4	NAG	L	1	1,4	14,14,15	0.35	0	17,19,21	0.62	0
4	NAG	L	2	4	14,14,15	0.27	0	17,19,21	0.40	0
4	NAG	M	1	1,4	14,14,15	0.35	0	17,19,21	0.39	0
4	NAG	M	2	4	14,14,15	0.18	0	17,19,21	0.39	0
5	NAG	N	1	1,5	14,14,15	0.35	0	17,19,21	0.41	0
5	NAG	N	2	5	14,14,15	0.21	0	17,19,21	0.46	0
5	BMA	N	3	5	11,11,12	0.62	0	15,15,17	0.87	0
4	NAG	O	1	2,4	14,14,15	0.35	0	17,19,21	0.43	0
4	NAG	O	2	4	14,14,15	0.22	0	17,19,21	0.42	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	J	1	1,4	-	2/6/23/26	0/1/1/1
4	NAG	J	2	4	-	2/6/23/26	0/1/1/1
4	NAG	K	1	1,4	-	2/6/23/26	0/1/1/1
4	NAG	K	2	4	-	2/6/23/26	0/1/1/1
4	NAG	L	1	1,4	-	4/6/23/26	0/1/1/1
4	NAG	L	2	4	-	1/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	M	1	1,4	-	2/6/23/26	0/1/1/1
4	NAG	M	2	4	-	2/6/23/26	0/1/1/1
5	NAG	N	1	1,5	-	0/6/23/26	0/1/1/1
5	NAG	N	2	5	-	2/6/23/26	0/1/1/1
5	BMA	N	3	5	-	1/2/19/22	0/1/1/1
4	NAG	O	1	2,4	-	0/6/23/26	0/1/1/1
4	NAG	O	2	4	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (22) torsion outliers are listed below:

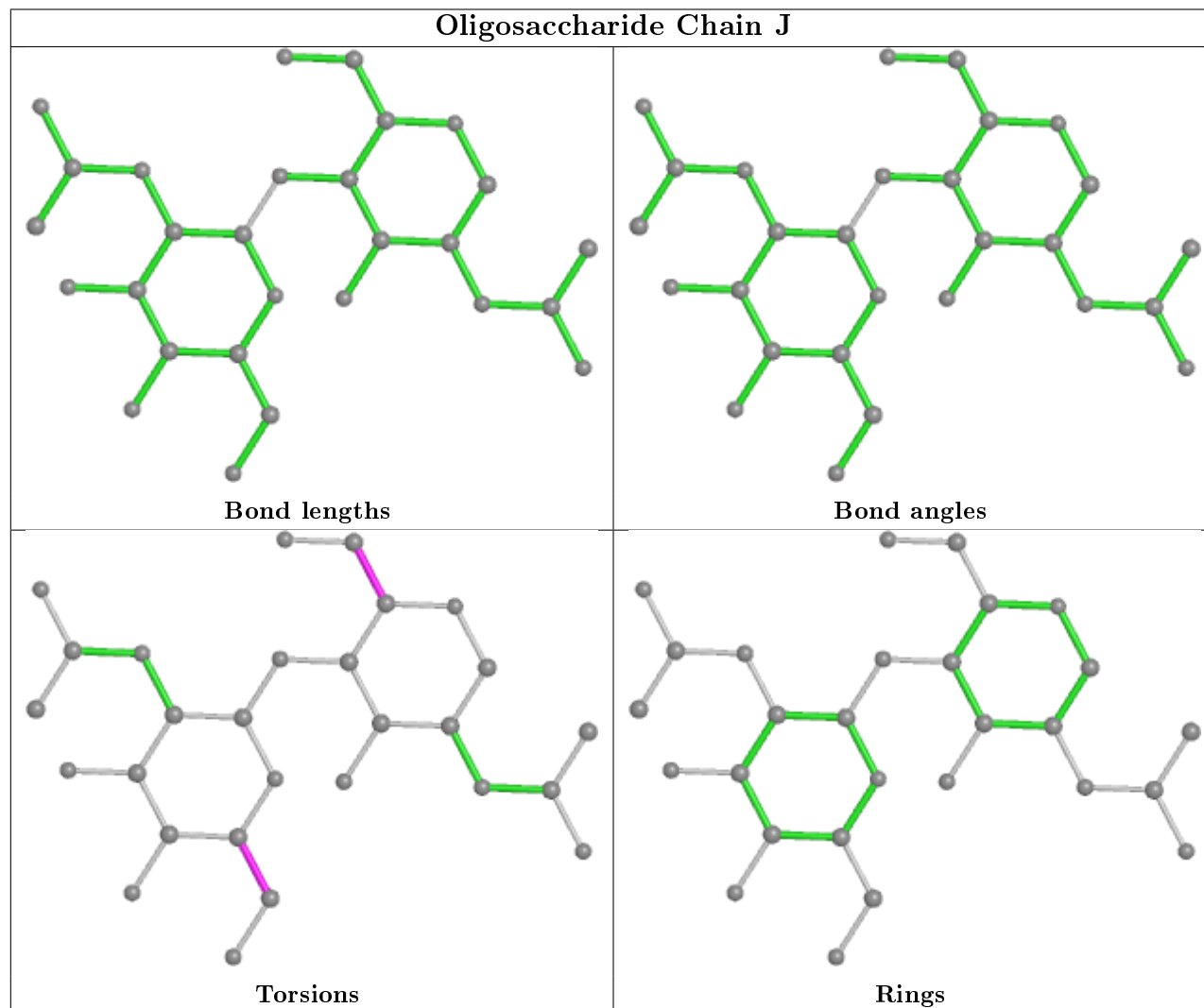
Mol	Chain	Res	Type	Atoms
4	M	1	NAG	O5-C5-C6-O6
4	L	1	NAG	O5-C5-C6-O6
5	N	2	NAG	O5-C5-C6-O6
4	O	2	NAG	O5-C5-C6-O6
4	O	2	NAG	C4-C5-C6-O6
4	J	1	NAG	O5-C5-C6-O6
4	M	1	NAG	C4-C5-C6-O6
4	L	1	NAG	C4-C5-C6-O6
5	N	2	NAG	C4-C5-C6-O6
4	J	1	NAG	C4-C5-C6-O6
4	K	2	NAG	O5-C5-C6-O6
4	J	2	NAG	O5-C5-C6-O6
4	M	2	NAG	O5-C5-C6-O6
4	M	2	NAG	C4-C5-C6-O6
4	K	1	NAG	O5-C5-C6-O6
4	K	1	NAG	C4-C5-C6-O6
4	K	2	NAG	C4-C5-C6-O6
4	J	2	NAG	C4-C5-C6-O6
4	L	2	NAG	O5-C5-C6-O6
4	L	1	NAG	C3-C2-N2-C7
5	N	3	BMA	C4-C5-C6-O6
4	L	1	NAG	C1-C2-N2-C7

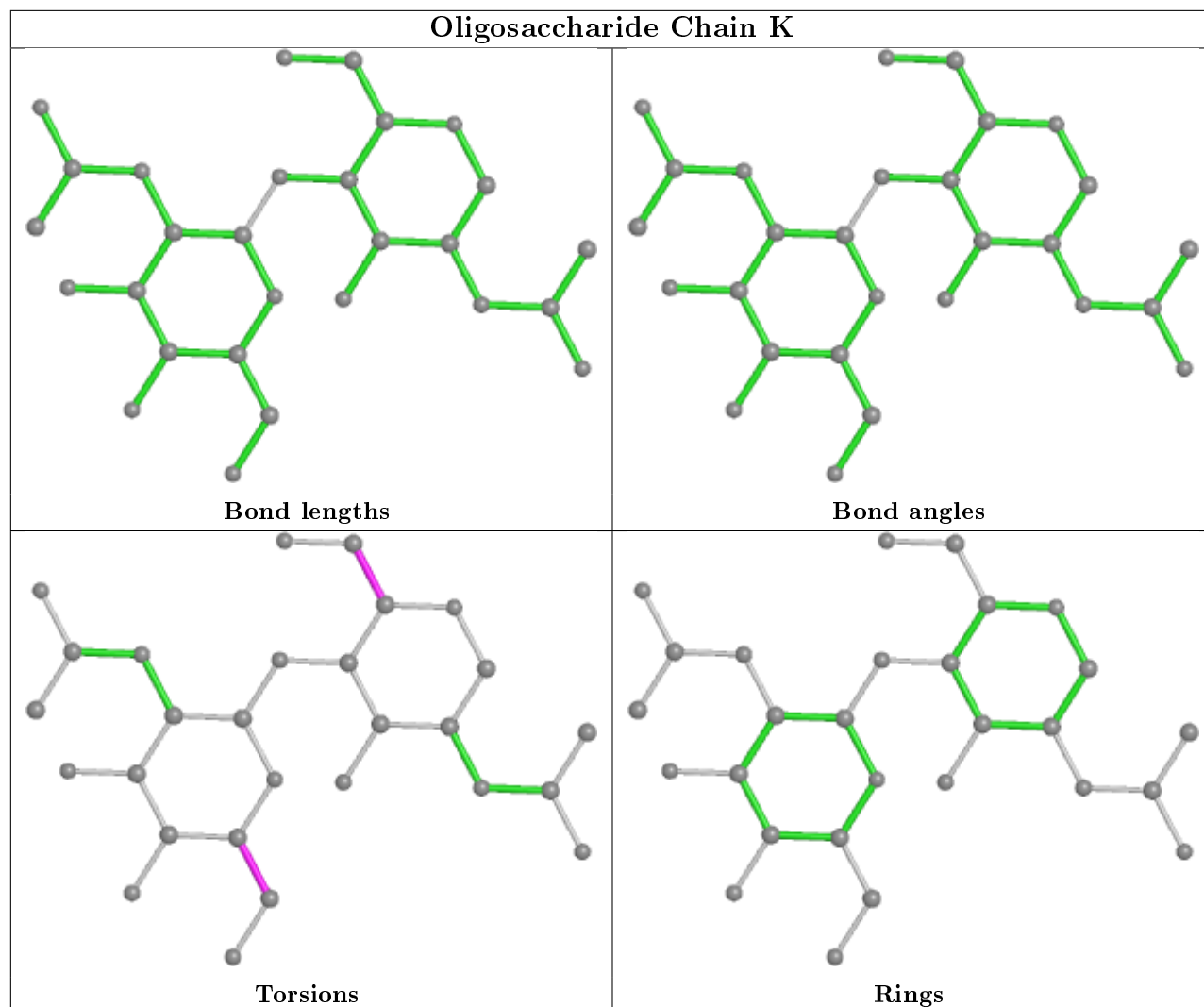
There are no ring outliers.

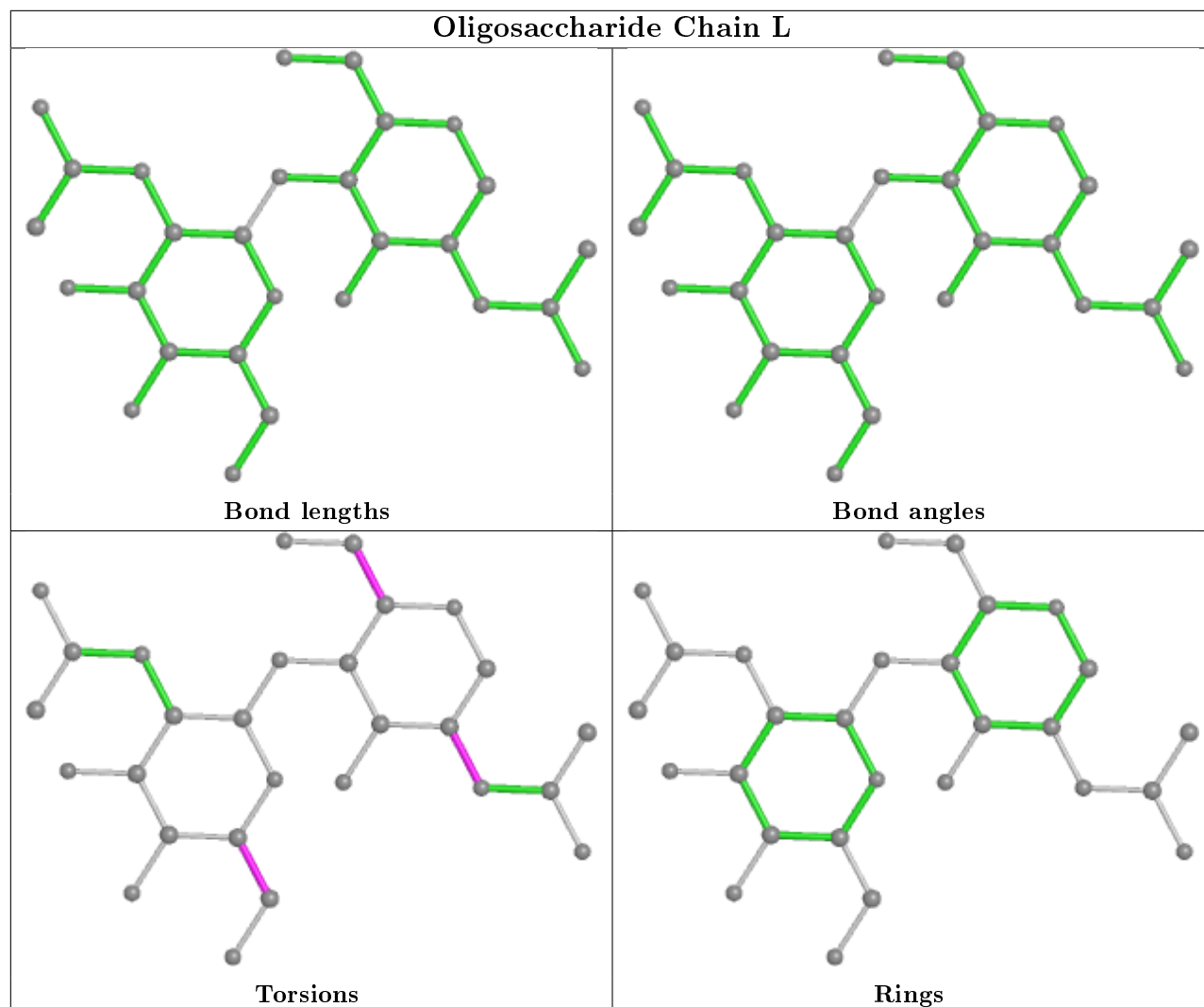
1 monomer is involved in 1 short contact:

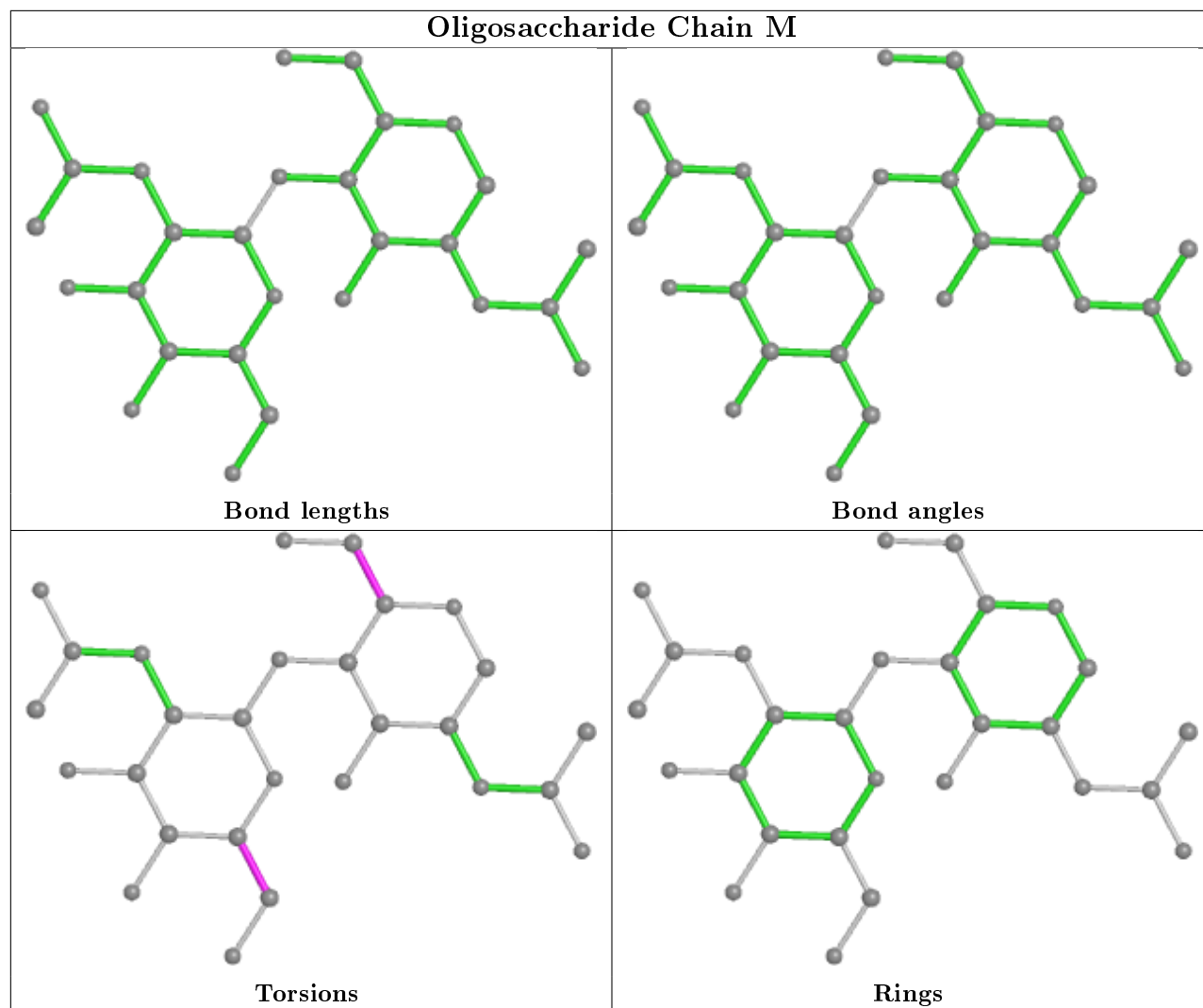
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	N	1	NAG	1	0

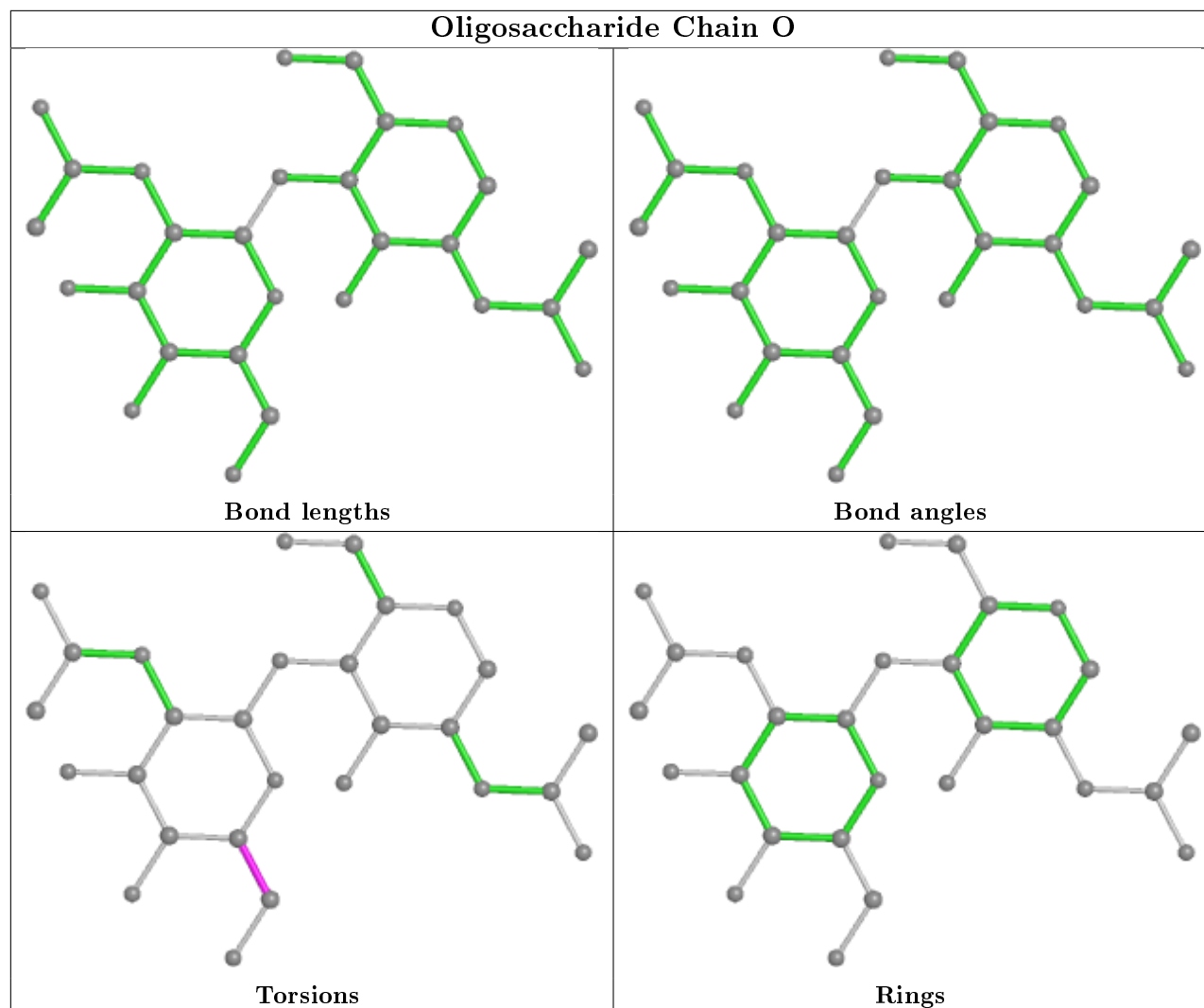
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

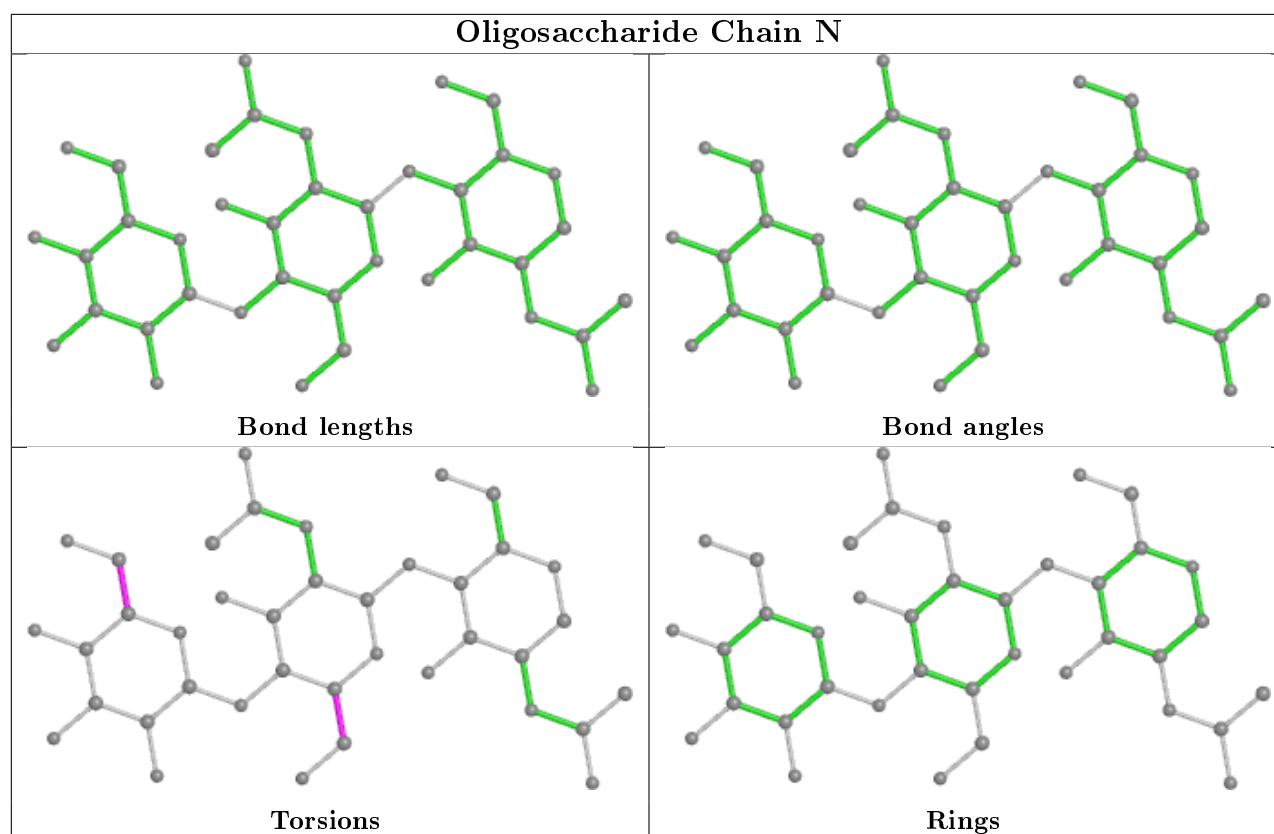












## 5.6 Ligand geometry [i](#)

6 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	A	402	1	14,14,15	0.23	0	17,19,21	0.31	0
6	NAG	E	405	1	14,14,15	0.28	0	17,19,21	0.42	0
6	NAG	D	601	2	14,14,15	0.47	0	17,19,21	0.46	0
6	NAG	E	401	1	14,14,15	0.29	0	17,19,21	0.52	0
6	NAG	B	601	2	14,14,15	0.37	0	17,19,21	0.47	0
6	NAG	A	401	1	14,14,15	0.36	0	17,19,21	0.40	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the



Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	A	402	1	-	2/6/23/26	0/1/1/1
6	NAG	E	405	1	-	3/6/23/26	0/1/1/1
6	NAG	D	601	2	-	2/6/23/26	0/1/1/1
6	NAG	E	401	1	-	0/6/23/26	0/1/1/1
6	NAG	B	601	2	-	0/6/23/26	0/1/1/1
6	NAG	A	401	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	402	NAG	C4-C5-C6-O6
6	D	601	NAG	C4-C5-C6-O6
6	A	402	NAG	O5-C5-C6-O6
6	D	601	NAG	O5-C5-C6-O6
6	E	405	NAG	C3-C2-N2-C7
6	E	405	NAG	C1-C2-N2-C7
6	E	405	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<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	319/323 (98%)	-0.26	0 100 100	28, 41, 56, 91	0
1	C	316/323 (97%)	-0.23	1 (0%) 94 93	34, 46, 66, 103	0
1	E	317/323 (98%)	-0.22	1 (0%) 94 93	30, 50, 78, 103	0
2	B	176/177 (99%)	-0.11	0 100 100	30, 42, 62, 76	0
2	D	172/177 (97%)	0.11	2 (1%) 79 76	30, 49, 75, 93	0
2	F	171/177 (96%)	-0.13	0 100 100	31, 41, 59, 64	0
3	G	84/97 (86%)	-0.08	1 (1%) 79 76	37, 47, 72, 81	0
3	H	85/97 (87%)	-0.00	2 (2%) 59 53	35, 47, 65, 82	0
3	I	53/97 (54%)	3.07	34 (64%) 0 0	80, 120, 143, 145	0
All	All	1693/1791 (94%)	-0.05	41 (2%) 59 53	28, 45, 78, 145	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	52	ILE	6.2
3	I	53	THR	5.9
3	I	35	LEU	5.9
3	I	33	ALA	5.6
3	I	19	LEU	5.5
3	I	38	VAL	5.5
3	I	50	TYR	5.4
3	I	26	ILE	5.3
3	I	54	LEU	5.1
3	I	37	PHE	5.0
3	I	70	ILE	4.9
3	I	20	ALA	4.6
3	I	49	TYR	4.5
3	I	48	ASN	4.5
3	I	36	THR	4.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	I	39	GLU	4.3
3	I	27	TYR	4.2
3	I	81	VAL	4.1
3	I	55	ALA	4.1
3	I	29	TYR	4.0
3	I	51	TYR	3.9
1	E	222	TRP	3.9
3	I	84	PHE	3.7
3	G	88	GLY	3.6
3	I	71	GLY	3.6
3	I	16	TYR	3.5
3	I	28	ASN	3.5
3	I	34	HIS	3.4
3	I	56	ALA	3.2
3	I	30	HIS	3.1
3	I	18	GLN	2.9
2	D	358	SER	2.8
3	I	23	ALA	2.8
3	I	25	ALA	2.6
2	D	472	LYS	2.4
3	H	88	GLY	2.3
3	I	17	GLN	2.3
3	I	24	VAL	2.3
1	C	174	PHE	2.2
3	I	21	ARG	2.1
3	H	8	VAL	2.1

## 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.

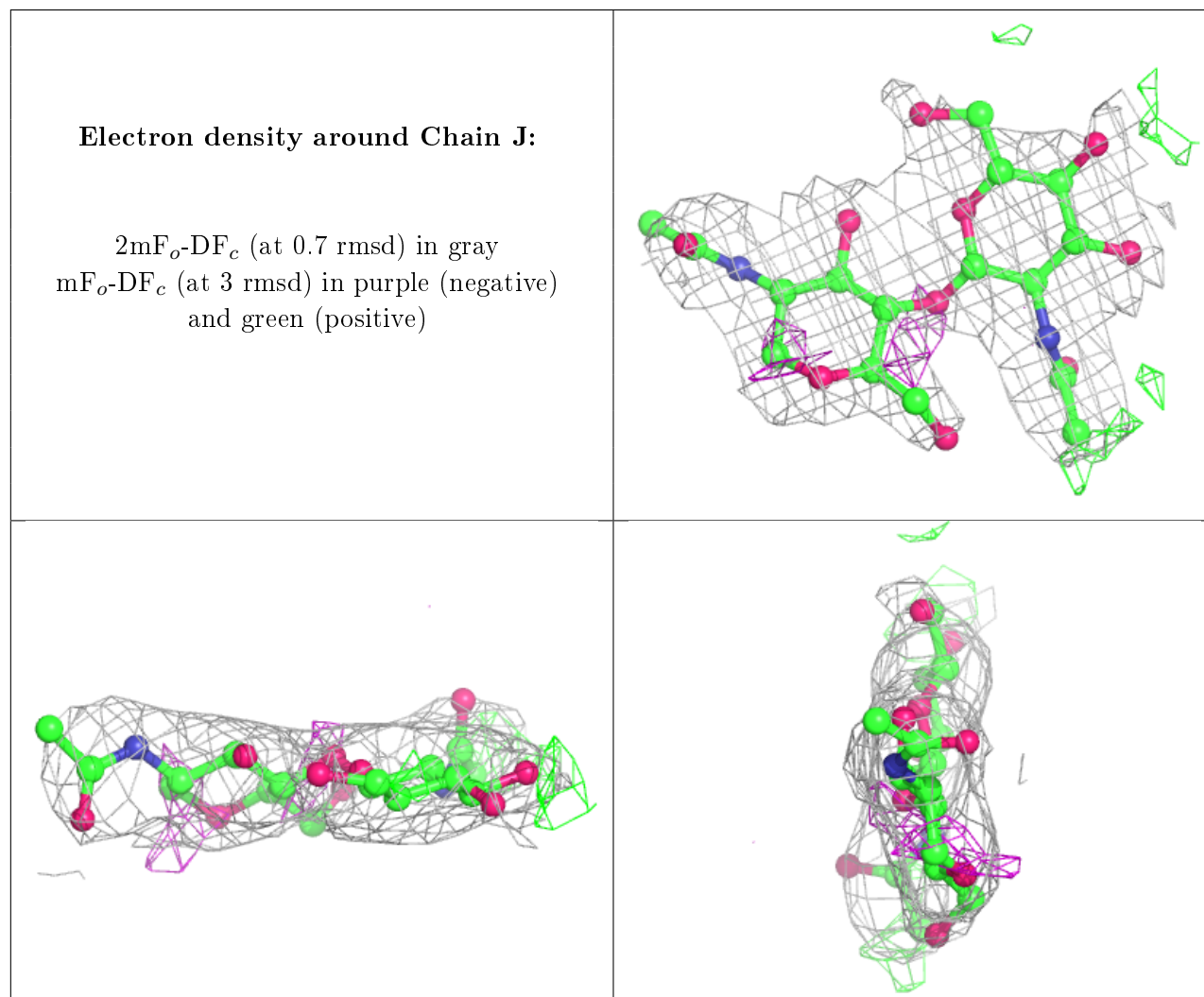
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	L	2	14/15	0.77	0.31	109,113,115,115	0
5	BMA	N	3	11/12	0.81	0.17	75,83,88,92	0

*Continued on next page...*

*Continued from previous page...*

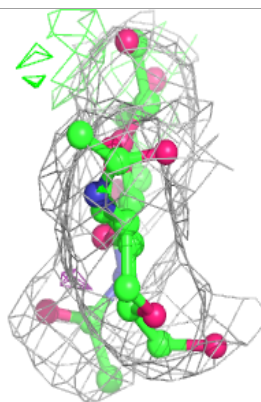
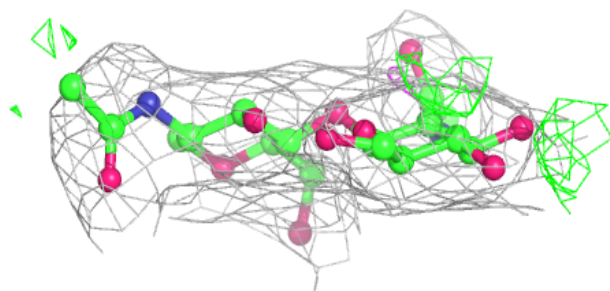
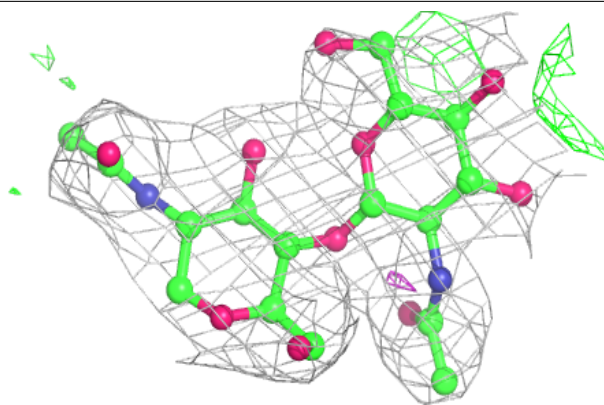
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	J	2	14/15	0.87	0.30	63,87,93,93	0
4	NAG	L	1	14/15	0.88	0.28	74,93,103,111	0
4	NAG	M	2	14/15	0.88	0.28	86,92,103,109	0
4	NAG	O	2	14/15	0.88	0.34	73,103,108,112	0
4	NAG	J	1	14/15	0.91	0.26	60,72,80,85	0
4	NAG	O	1	14/15	0.91	0.23	51,72,80,87	0
4	NAG	K	2	14/15	0.93	0.17	40,73,81,82	0
5	NAG	N	1	14/15	0.94	0.14	55,66,80,87	0
4	NAG	M	1	14/15	0.95	0.18	43,57,60,73	0
4	NAG	K	1	14/15	0.96	0.12	42,61,68,72	0
5	NAG	N	2	14/15	0.96	0.11	45,58,75,76	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.



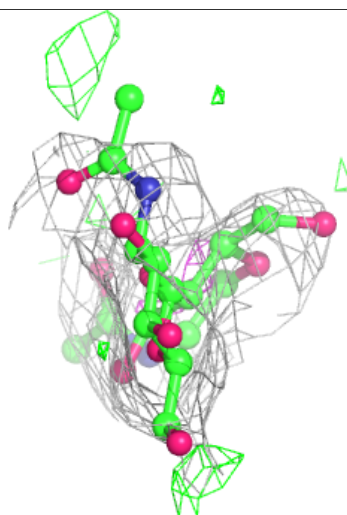
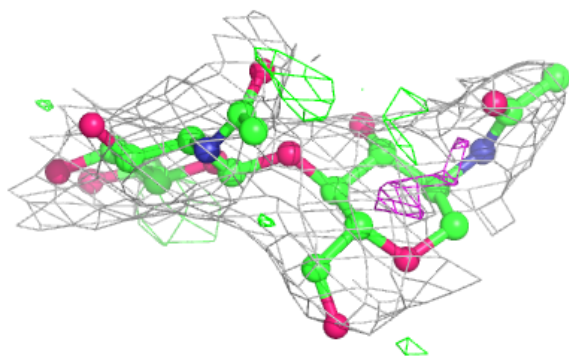
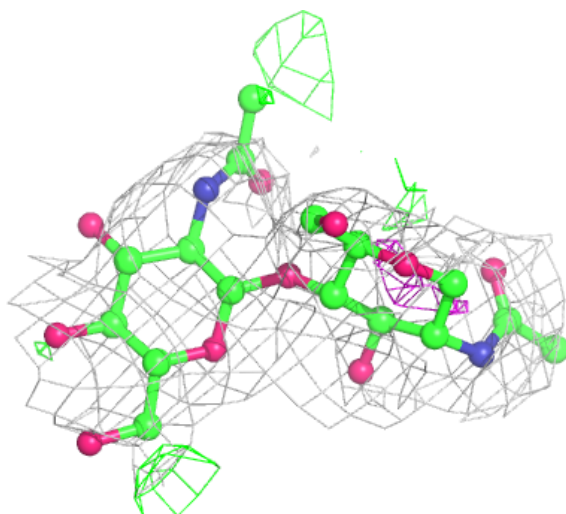
**Electron density around Chain K:**

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



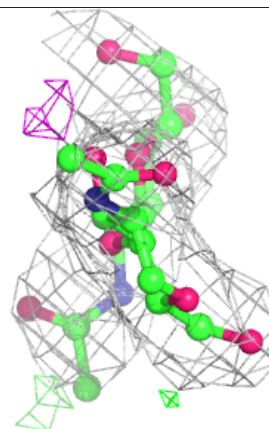
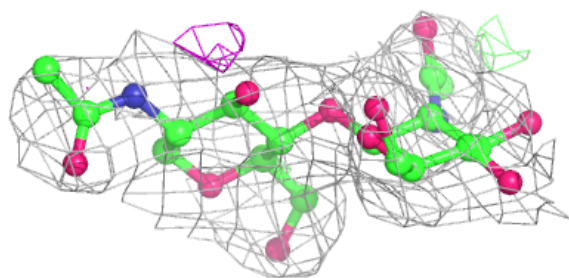
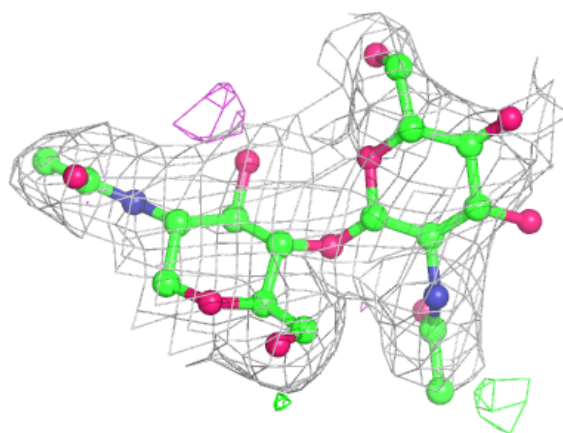
**Electron density around Chain L:**

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



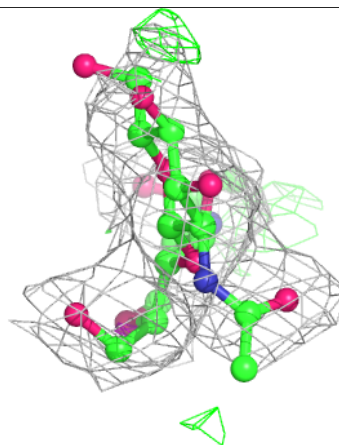
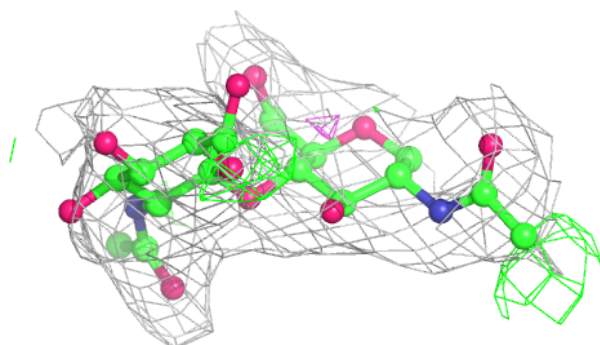
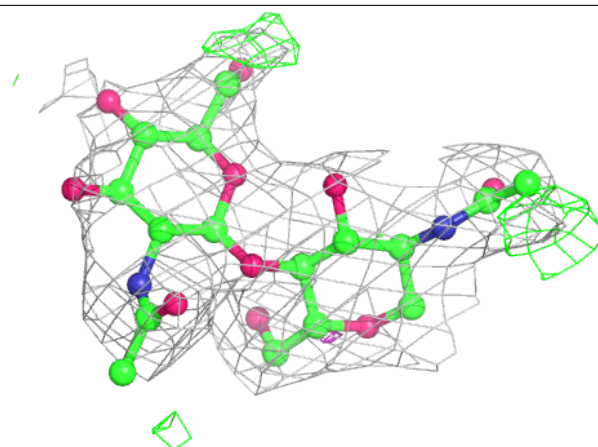
**Electron density around Chain M:**

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

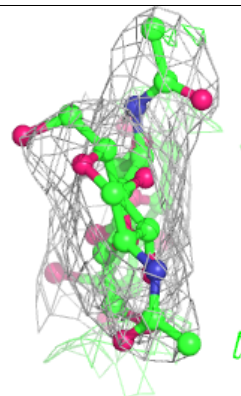
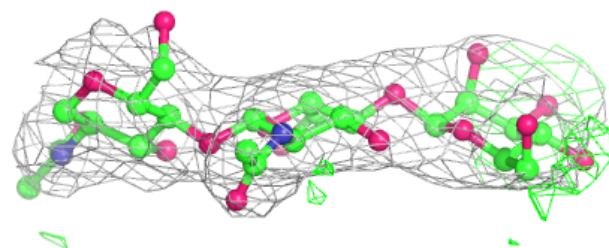
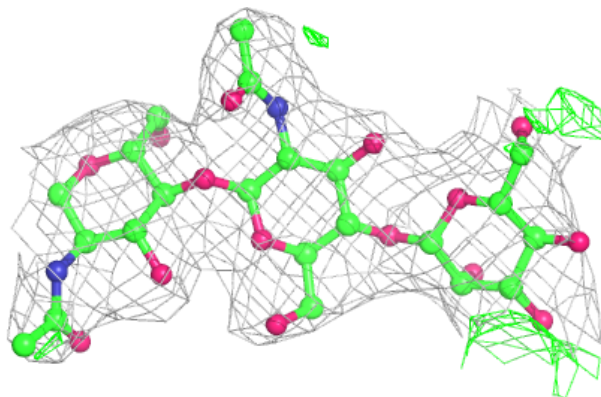


**Electron density around Chain O:**

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

**Electron density around Chain 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	405	14/15	0.81	0.28	47,80,94,96	0
6	NAG	A	402	14/15	0.86	0.29	58,75,84,84	0
6	NAG	D	601	14/15	0.88	0.24	60,75,85,91	0
6	NAG	B	601	14/15	0.88	0.23	60,65,81,84	0
6	NAG	E	401	14/15	0.89	0.21	46,60,66,66	0
6	NAG	A	401	14/15	0.90	0.18	43,58,67,71	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.