



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 24, 2025 – 05:26 am BST

PDB ID : 6TVF / pdb\_00006tvf  
Title : Crystal structure of the haemagglutinin from a H10N7 seal influenza virus isolated in Germany in complex with human receptor analogue, 6'-SLN  
Authors : Zhang, J.; Xiong, X.; Purkiss, A.; Walker, P.; Gamblin, S.; Skehel, J.J.  
Deposited on : 2020-01-09  
Resolution : 2.60 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 : 4-5-2 with Phenix2.0  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 2.0  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.46

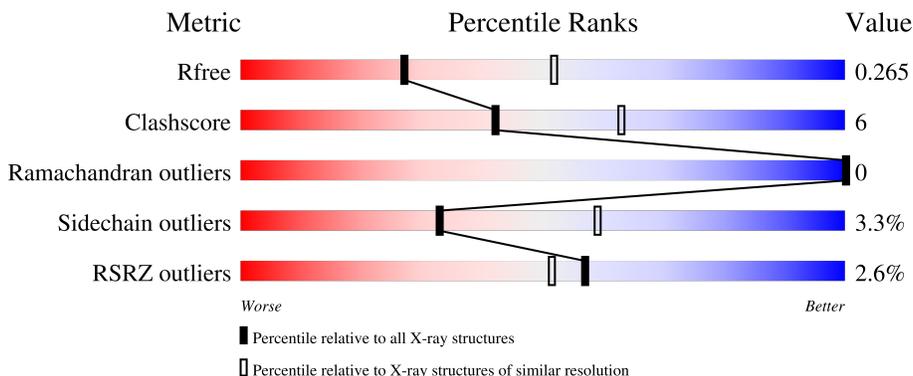
# 1 Overall quality at a glance

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.



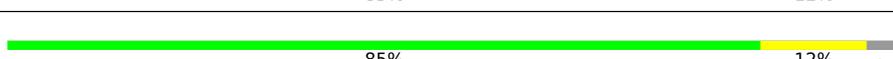
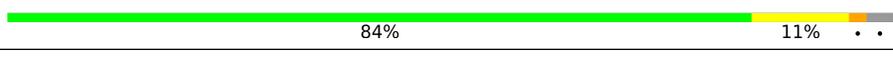
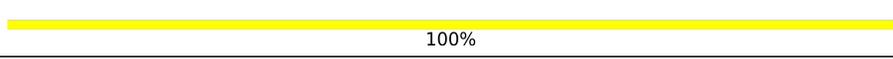
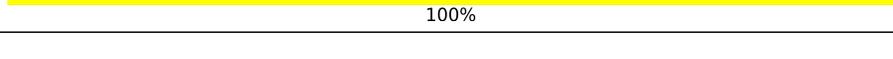
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	164625	3775 (2.60-2.60)
Clashscore	180529	4181 (2.60-2.60)
Ramachandran outliers	177936	4129 (2.60-2.60)
Sidechain outliers	177891	4129 (2.60-2.60)
RSRZ outliers	164620	3775 (2.60-2.60)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	325	 89% 8% ..
1	C	325	 2% 90% 8% ..
1	E	325	 2% 85% 12% ..
1	G	325	 14% 78% 19% ..
1	I	325	 4% 83% 14% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	325	 87% 10% ..
2	B	177	 85% 12% .
2	D	177	 86% 11% ..
2	F	177	 85% 12% .
2	H	177	 85% 12% .
2	J	177	 86% 11% .
2	L	177	 84% 11% ..
3	M	3	 67% 33%
3	P	3	 100%
3	R	3	 33% 67%
3	U	3	 67% 33%
4	N	2	 50% 50%
5	O	2	 100%
5	Q	2	 100%
5	S	2	 100%
5	T	2	 100%
5	V	2	 50% 50%

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 23705 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Hemagglutinin HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	320	2448	1517	443	472	16	0	1	0
1	C	320	2442	1514	440	472	16	0	1	0
1	E	320	2442	1514	440	472	16	0	1	0
1	G	320	2410	1489	439	466	16	0	0	0
1	I	320	2442	1514	440	472	16	0	1	0
1	K	318	2428	1505	441	466	16	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ASP	-	expression tag	UNP A0A0A7HR51
A	0	PRO	-	expression tag	UNP A0A0A7HR51
A	219	GLN	LEU	conflict	UNP A0A0A7HR51
C	-1	ASP	-	expression tag	UNP A0A0A7HR51
C	0	PRO	-	expression tag	UNP A0A0A7HR51
C	219	GLN	LEU	conflict	UNP A0A0A7HR51
E	-1	ASP	-	expression tag	UNP A0A0A7HR51
E	0	PRO	-	expression tag	UNP A0A0A7HR51
E	219	GLN	LEU	conflict	UNP A0A0A7HR51
G	-1	ASP	-	expression tag	UNP A0A0A7HR51
G	0	PRO	-	expression tag	UNP A0A0A7HR51
G	219	GLN	LEU	conflict	UNP A0A0A7HR51
I	-1	ASP	-	expression tag	UNP A0A0A7HR51
I	0	PRO	-	expression tag	UNP A0A0A7HR51
I	219	GLN	LEU	conflict	UNP A0A0A7HR51
K	-1	ASP	-	expression tag	UNP A0A0A7HR51
K	0	PRO	-	expression tag	UNP A0A0A7HR51

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
K	219	GLN	LEU	conflict	UNP A0A0A7HR51

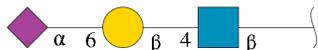
- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	172	Total	C	N	O	S	0	0	0
			1386	857	241	280	8			
2	D	172	Total	C	N	O	S	0	0	0
			1383	855	240	280	8			
2	F	172	Total	C	N	O	S	0	0	0
			1383	855	240	280	8			
2	H	172	Total	C	N	O	S	0	0	0
			1383	855	240	280	8			
2	J	172	Total	C	N	O	S	0	0	0
			1383	855	240	280	8			
2	L	172	Total	C	N	O	S	0	0	0
			1383	855	240	280	8			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	177	LYS	-	expression tag	UNP A0A0A7HR51
D	177	LYS	-	expression tag	UNP A0A0A7HR51
F	177	LYS	-	expression tag	UNP A0A0A7HR51
H	177	LYS	-	expression tag	UNP A0A0A7HR51
J	177	LYS	-	expression tag	UNP A0A0A7HR51
L	177	LYS	-	expression tag	UNP A0A0A7HR51

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



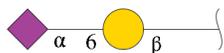
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	M	3	Total	C	N	O	0	0	0
			46	25	2	19			
3	P	3	Total	C	N	O	0	0	0
			46	25	2	19			
3	R	3	Total	C	N	O	0	0	0
			46	25	2	19			

Continued on next page...

Continued from previous page...

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

- Molecule 4 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	N	2	32	17	1	14	0	0	0

- Molecule 5 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
			Total	C	N	O			
5	O	2	28	16	2	10	0	0	0
5	Q	2	28	16	2	10	0	0	0
5	S	2	28	16	2	10	0	0	0
5	T	2	28	16	2	10	0	0	0
5	V	2	28	16	2	10	0	0	0

- Molecule 6 is CALCIUM ION (CCD ID: CA) (formula: Ca).

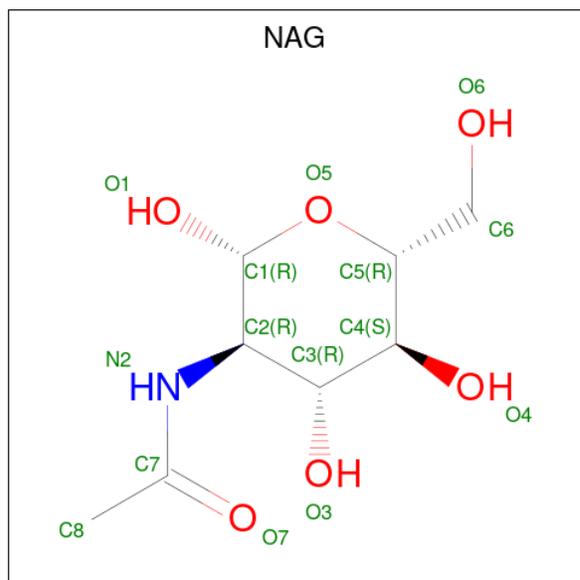
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Ca	0	0
			1	1		
6	I	1	Total	Ca	0	0
			1	1		

Continued on next page...

Continued from previous page...

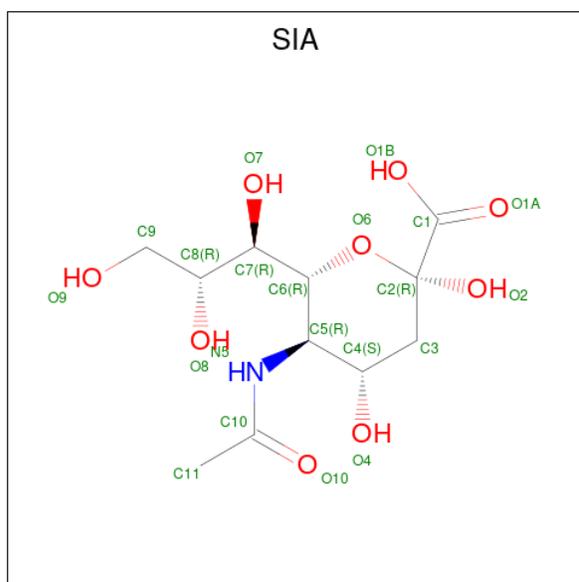
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	K	1	Total Ca 1 1	0	0

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula:  $C_8H_{15}NO_6$ ) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 8 is N-acetyl-alpha-neuraminic acid (CCD ID: SIA) (formula:  $C_{11}H_{19}NO_9$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	N			O
8	I	1	21	11	1	9	0	0

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	34	Total	O	0	0
			34	34		
9	B	22	Total	O	0	0
			22	22		
9	C	45	Total	O	0	0
			45	45		
9	D	37	Total	O	0	0
			37	37		
9	E	24	Total	O	0	0
			24	24		
9	F	28	Total	O	0	0
			28	28		
9	G	12	Total	O	0	0
			12	12		
9	H	28	Total	O	0	0
			28	28		
9	I	7	Total	O	0	0
			7	7		
9	J	20	Total	O	0	0
			20	20		
9	K	79	Total	O	0	0
			79	79		

*Continued on next page...*

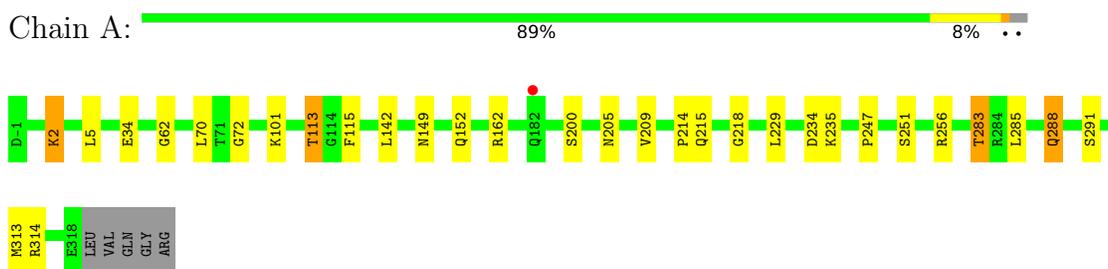
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
9	L	34	Total	O	0	0
			34	34		

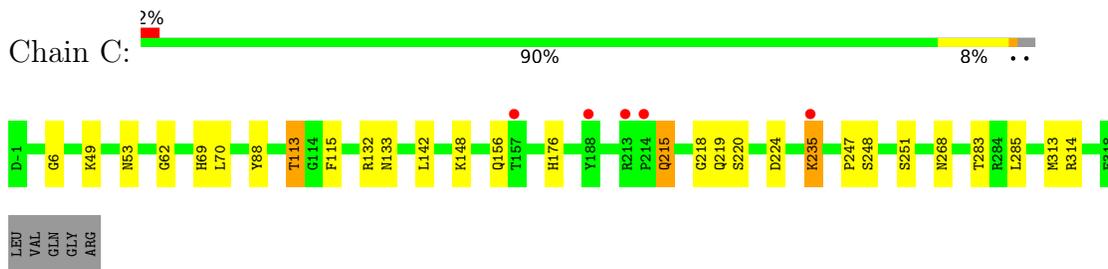
### 3 Residue-property plots

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

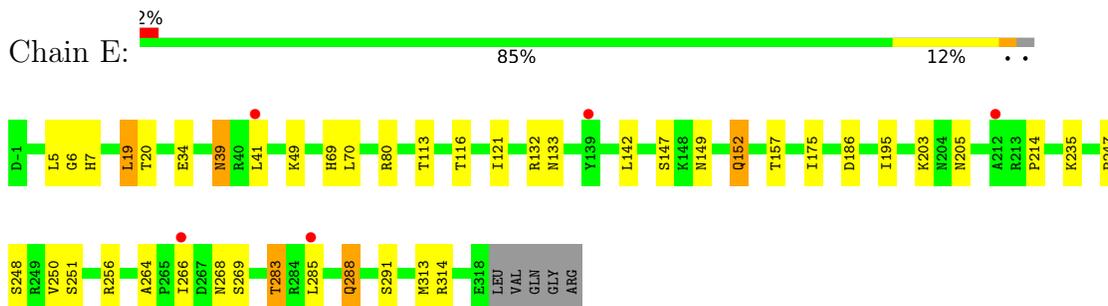
- Molecule 1: Hemagglutinin HA1



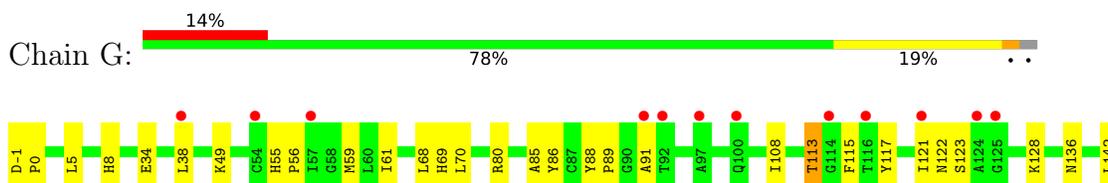
- Molecule 1: Hemagglutinin HA1

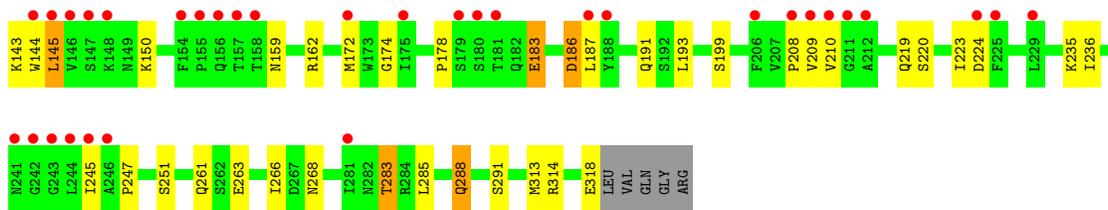


- Molecule 1: Hemagglutinin HA1

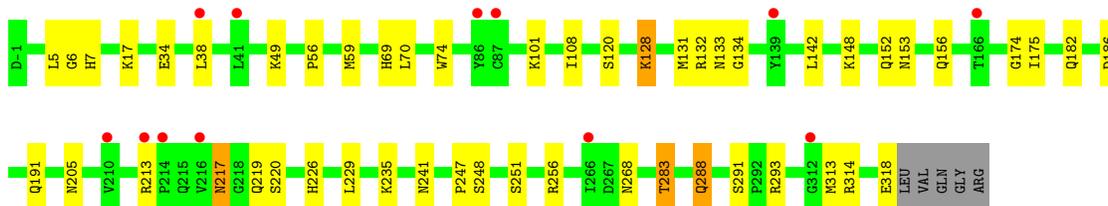
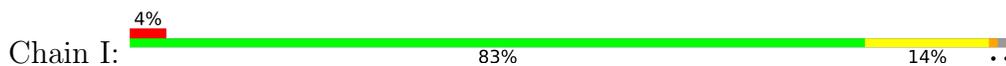


- Molecule 1: Hemagglutinin HA1

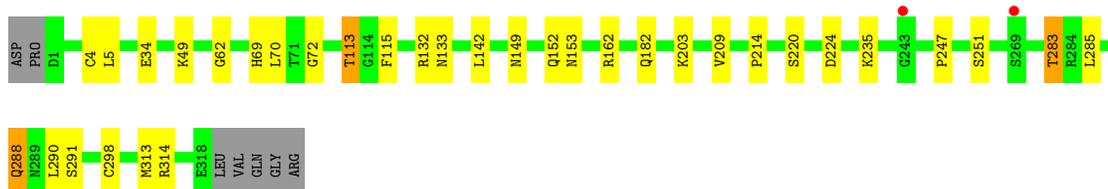
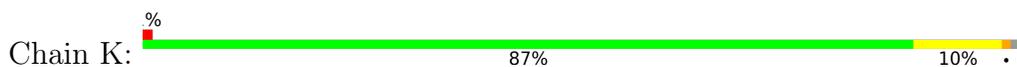




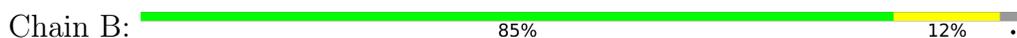
- Molecule 1: Hemagglutinin HA1



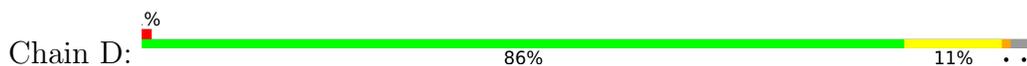
- Molecule 1: Hemagglutinin HA1



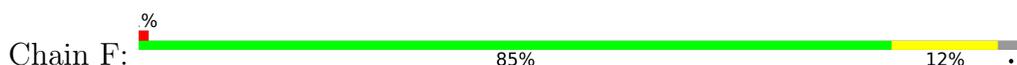
- Molecule 2: Hemagglutinin HA2



- Molecule 2: Hemagglutinin HA2



- Molecule 2: Hemagglutinin HA2



- Molecule 2: Hemagglutinin HA2

Chain H:  85% 12%



- Molecule 2: Hemagglutinin HA2

Chain J:  2% 86% 11%



- Molecule 2: Hemagglutinin HA2

Chain L:  84% 11%

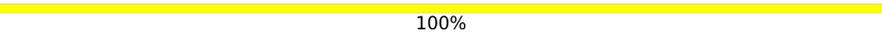


- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain M:  67% 33%



- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain P:  100%



- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain R:  33% 67%



- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain U:  67% 33%

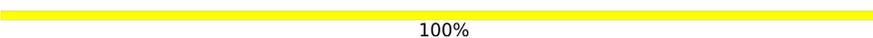
MAG1  
GAL2  
SIA3

- Molecule 4: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose

Chain N:  50% 50%

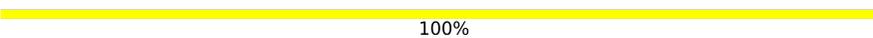
GAL1  
SIA2

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

Chain O:  100%

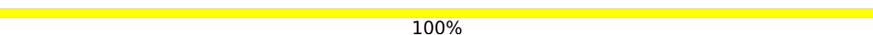
MAG1  
MAG2

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

Chain Q:  100%

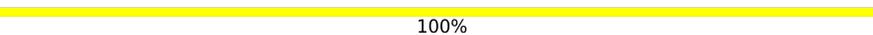
MAG1  
MAG2

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

Chain S:  100%

MAG1  
MAG2

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

Chain T:  100%

MAG1  
MAG2

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

Chain V:  50% 50%

MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.81Å 214.14Å 158.39Å 90.00° 101.07° 90.00°	Depositor
Resolution (Å)	68.52 – 2.60 68.52 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.0 (68.52-2.60) 98.9 (68.52-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.64 (at 2.61Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.239 , 0.276 0.233 , 0.265	Depositor DCC
$R_{free}$ test set	7071 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	59.1	Xtrriage
Anisotropy	0.538	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 45.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	23705	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	91.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 30.97 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1961e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: GAL, SIA, NAG, CA

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	1.01	0/2501	1.30	1/3389 (0.0%)
1	C	1.05	0/2495	1.31	1/3382 (0.0%)
1	E	1.02	0/2495	1.33	1/3382 (0.0%)
1	G	1.05	0/2456	1.35	2/3327 (0.1%)
1	I	1.05	1/2495 (0.0%)	1.31	2/3382 (0.1%)
1	K	1.02	0/2477	1.31	1/3355 (0.0%)
2	B	0.96	0/1411	1.48	0/1903
2	D	0.96	0/1408	1.49	0/1900
2	F	0.96	0/1408	1.47	0/1900
2	H	0.95	0/1408	1.49	0/1900
2	J	0.96	0/1408	1.48	0/1900
2	L	0.97	0/1408	1.48	0/1900
All	All	1.01	1/23370 (0.0%)	1.38	8/31620 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	74	TRP	C-O	-5.01	1.17	1.23

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	186	ASP	CA-CB-CG	7.60	120.20	112.60
1	K	72	GLY	CA-C-O	-5.72	117.47	122.16
1	G	191	GLN	CA-C-O	-5.64	115.41	121.56
1	C	156	GLN	CA-C-O	-5.63	114.55	121.02
1	G	186	ASP	CA-CB-CG	5.49	118.09	112.60

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	2448	0	2400	24	0
1	C	2442	0	2389	19	0
1	E	2442	0	2389	34	0
1	G	2410	0	2352	54	0
1	I	2442	0	2389	37	0
1	K	2428	0	2385	27	0
2	B	1386	0	1290	22	0
2	D	1383	0	1282	21	0
2	F	1383	0	1282	30	0
2	H	1383	0	1283	18	0
2	J	1383	0	1282	24	0
2	L	1383	0	1281	22	0
3	M	46	0	40	1	0
3	P	46	0	40	0	0
3	R	46	0	40	4	0
3	U	46	0	40	1	0
4	N	32	0	28	1	0
5	O	28	0	25	0	0
5	Q	28	0	25	0	0
5	S	28	0	25	0	0
5	T	28	0	25	0	0
5	V	28	0	25	0	0
6	A	1	0	0	0	0
6	I	1	0	0	0	0
6	K	1	0	0	0	0
7	B	28	0	26	0	0
7	L	14	0	13	1	0
8	I	21	0	18	2	0
9	A	34	0	0	1	0
9	B	22	0	0	2	0
9	C	45	0	0	2	0
9	D	37	0	0	4	0
9	E	24	0	0	1	0
9	F	28	0	0	0	0
9	G	12	0	0	0	0
9	H	28	0	0	1	0
9	I	7	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	J	20	0	0	2	0
9	K	79	0	0	3	0
9	L	34	0	0	0	0
All	All	23705	0	22374	281	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 6.

The worst 5 of 281 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:95:GLN:HE22	2:J:95:GLN:HE21	1.14	0.90
2:F:95:GLN:HE21	2:L:95:GLN:HE21	1.20	0.88
2:F:25:ARG:HE	2:F:34:GLN:HE21	1.22	0.87
1:I:34:GLU:OE1	1:I:283:THR:HG21	1.76	0.85
2:B:95:GLN:HE21	2:J:95:GLN:HE21	1.22	0.84

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	319/325 (98%)	315 (99%)	4 (1%)	0	100	100
1	C	319/325 (98%)	317 (99%)	2 (1%)	0	100	100
1	E	319/325 (98%)	316 (99%)	3 (1%)	0	100	100
1	G	318/325 (98%)	316 (99%)	2 (1%)	0	100	100
1	I	319/325 (98%)	317 (99%)	2 (1%)	0	100	100
1	K	316/325 (97%)	311 (98%)	5 (2%)	0	100	100
2	B	170/177 (96%)	167 (98%)	3 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	170/177 (96%)	166 (98%)	4 (2%)	0	100	100
2	F	170/177 (96%)	167 (98%)	3 (2%)	0	100	100
2	H	170/177 (96%)	166 (98%)	4 (2%)	0	100	100
2	J	170/177 (96%)	167 (98%)	3 (2%)	0	100	100
2	L	170/177 (96%)	166 (98%)	4 (2%)	0	100	100
All	All	2930/3012 (97%)	2891 (99%)	39 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	271/275 (98%)	266 (98%)	5 (2%)	54	77
1	C	270/275 (98%)	262 (97%)	8 (3%)	36	63
1	E	270/275 (98%)	259 (96%)	11 (4%)	26	51
1	G	263/275 (96%)	246 (94%)	17 (6%)	14	31
1	I	270/275 (98%)	256 (95%)	14 (5%)	19	41
1	K	268/275 (98%)	263 (98%)	5 (2%)	52	75
2	B	146/151 (97%)	143 (98%)	3 (2%)	48	73
2	D	145/151 (96%)	141 (97%)	4 (3%)	38	65
2	F	145/151 (96%)	143 (99%)	2 (1%)	62	82
2	H	145/151 (96%)	142 (98%)	3 (2%)	48	73
2	J	145/151 (96%)	141 (97%)	4 (3%)	38	65
2	L	145/151 (96%)	140 (97%)	5 (3%)	32	58
All	All	2483/2556 (97%)	2402 (97%)	81 (3%)	33	59

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	213	ARG
1	K	113	THR
1	I	220	SER
1	I	293	ARG
1	K	288	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 118 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	161	GLN
2	L	30	GLN
2	H	34	GLN
2	L	27	GLN
1	K	23	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](#)

24 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	3	15,15,15	0.84	0	21,21,21	2.27	7 (33%)
3	GAL	M	2	3	11,11,12	0.43	0	15,15,17	1.34	2 (13%)
3	SIA	M	3	3	20,20,21	0.80	1 (5%)	24,28,31	1.71	7 (29%)
4	GAL	N	1	4	12,12,12	0.73	0	17,17,17	1.41	3 (17%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	SIA	N	2	4	20,20,21	0.65	0	24,28,31	1.33	3 (12%)
5	NAG	O	1	6,2,5	14,14,15	0.58	0	17,19,21	1.82	3 (17%)
5	NAG	O	2	5	14,14,15	0.46	0	17,19,21	1.21	1 (5%)
3	NAG	P	1	3	15,15,15	0.81	0	21,21,21	1.98	3 (14%)
3	GAL	P	2	3	11,11,12	0.62	0	15,15,17	1.62	2 (13%)
3	SIA	P	3	3	20,20,21	0.71	0	24,28,31	1.54	5 (20%)
5	NAG	Q	1	2,5	14,14,15	0.52	0	17,19,21	1.52	3 (17%)
5	NAG	Q	2	5	14,14,15	0.49	0	17,19,21	1.12	1 (5%)
3	NAG	R	1	3	15,15,15	0.51	0	21,21,21	2.18	8 (38%)
3	GAL	R	2	3	11,11,12	0.39	0	15,15,17	1.29	3 (20%)
3	SIA	R	3	3	20,20,21	0.78	0	24,28,31	1.42	6 (25%)
5	NAG	S	1	6,5	14,14,15	0.57	0	17,19,21	1.98	4 (23%)
5	NAG	S	2	5	14,14,15	0.40	0	17,19,21	0.90	1 (5%)
5	NAG	T	1	2,5	14,14,15	0.43	0	17,19,21	1.66	1 (5%)
5	NAG	T	2	5	14,14,15	0.35	0	17,19,21	1.35	1 (5%)
3	NAG	U	1	3	15,15,15	0.82	0	21,21,21	1.68	4 (19%)
3	GAL	U	2	3	11,11,12	0.28	0	15,15,17	1.49	3 (20%)
3	SIA	U	3	3	20,20,21	0.71	0	24,28,31	1.73	6 (25%)
5	NAG	V	1	2,5	14,14,15	0.39	0	17,19,21	1.18	2 (11%)
5	NAG	V	2	5	14,14,15	0.52	0	17,19,21	0.87	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
3	NAG	M	1	3	-	6/6/26/26	0/1/1/1
3	GAL	M	2	3	-	0/2/19/22	0/1/1/1
3	SIA	M	3	3	-	2/18/34/38	0/1/1/1
4	GAL	N	1	4	-	2/2/22/22	0/1/1/1
4	SIA	N	2	4	-	0/18/34/38	0/1/1/1
5	NAG	O	1	6,2,5	-	0/6/23/26	0/1/1/1
5	NAG	O	2	5	-	0/6/23/26	0/1/1/1
3	NAG	P	1	3	-	2/6/26/26	0/1/1/1
3	GAL	P	2	3	-	1/2/19/22	0/1/1/1
3	SIA	P	3	3	-	0/18/34/38	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	Q	1	2,5	-	0/6/23/26	0/1/1/1
5	NAG	Q	2	5	-	2/6/23/26	0/1/1/1
3	NAG	R	1	3	-	4/6/26/26	0/1/1/1
3	GAL	R	2	3	-	0/2/19/22	0/1/1/1
3	SIA	R	3	3	-	1/18/34/38	0/1/1/1
5	NAG	S	1	6,5	-	1/6/23/26	0/1/1/1
5	NAG	S	2	5	-	2/6/23/26	0/1/1/1
5	NAG	T	1	2,5	-	2/6/23/26	0/1/1/1
5	NAG	T	2	5	-	2/6/23/26	0/1/1/1
3	NAG	U	1	3	-	3/6/26/26	0/1/1/1
3	GAL	U	2	3	-	0/2/19/22	0/1/1/1
3	SIA	U	3	3	-	0/18/34/38	0/1/1/1
5	NAG	V	1	2,5	-	0/6/23/26	0/1/1/1
5	NAG	V	2	5	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	M	3	SIA	C2-C1	2.00	1.54	1.52

The worst 5 of 79 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	T	1	NAG	C1-O5-C5	5.83	120.09	112.19
5	S	1	NAG	C1-O5-C5	5.64	119.84	112.19
3	P	1	NAG	O5-C1-C2	5.59	115.13	109.52
3	P	1	NAG	C1-C2-N2	-4.96	104.98	110.73
3	P	2	GAL	C1-C2-C3	4.93	115.72	109.67

There are no chirality outliers.

5 of 32 torsion outliers are listed below:

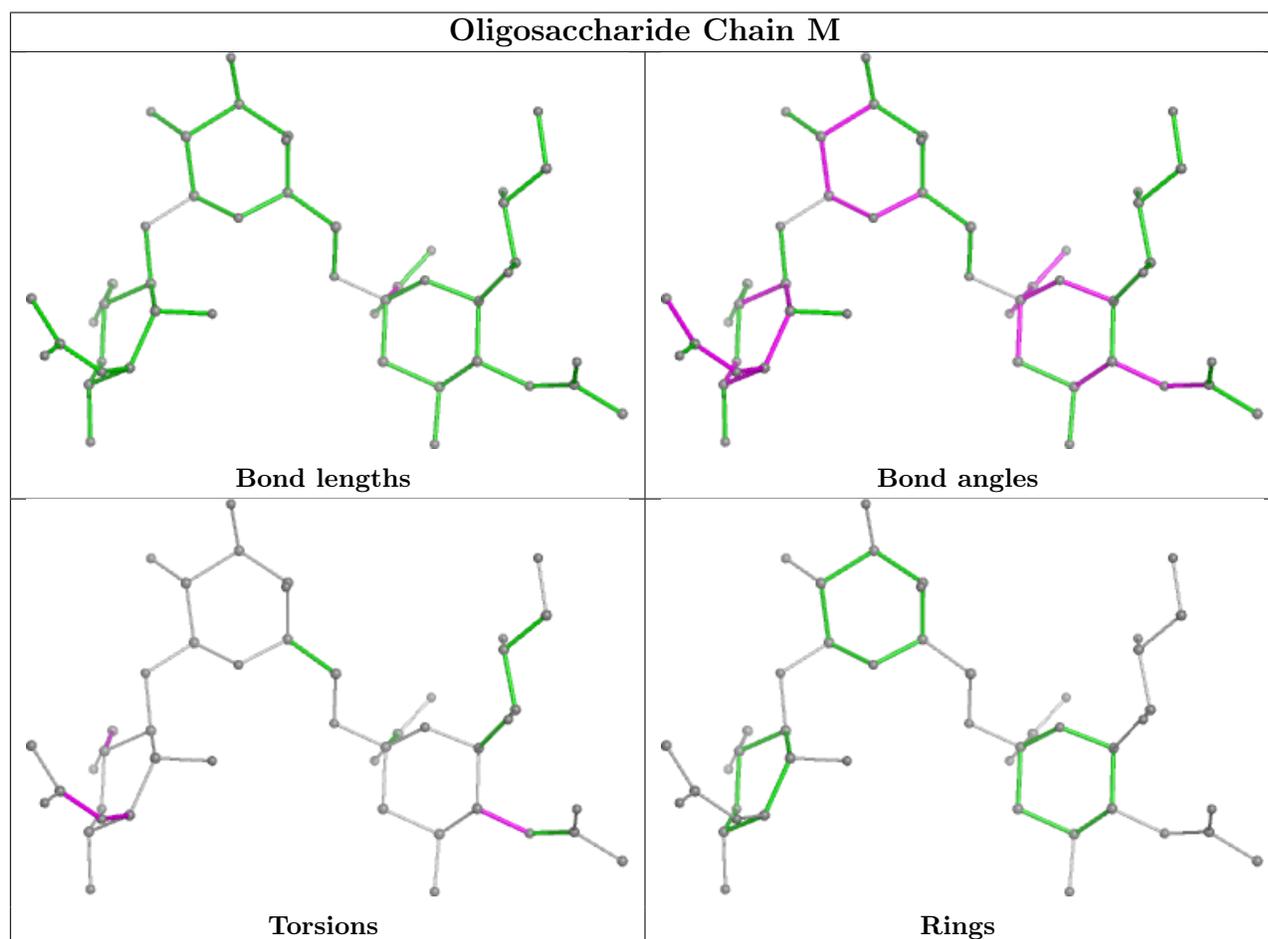
Mol	Chain	Res	Type	Atoms
3	P	1	NAG	C1-C2-N2-C7
4	N	1	GAL	O5-C5-C6-O6
5	T	1	NAG	O5-C5-C6-O6
3	M	1	NAG	O5-C5-C6-O6
5	Q	2	NAG	C4-C5-C6-O6

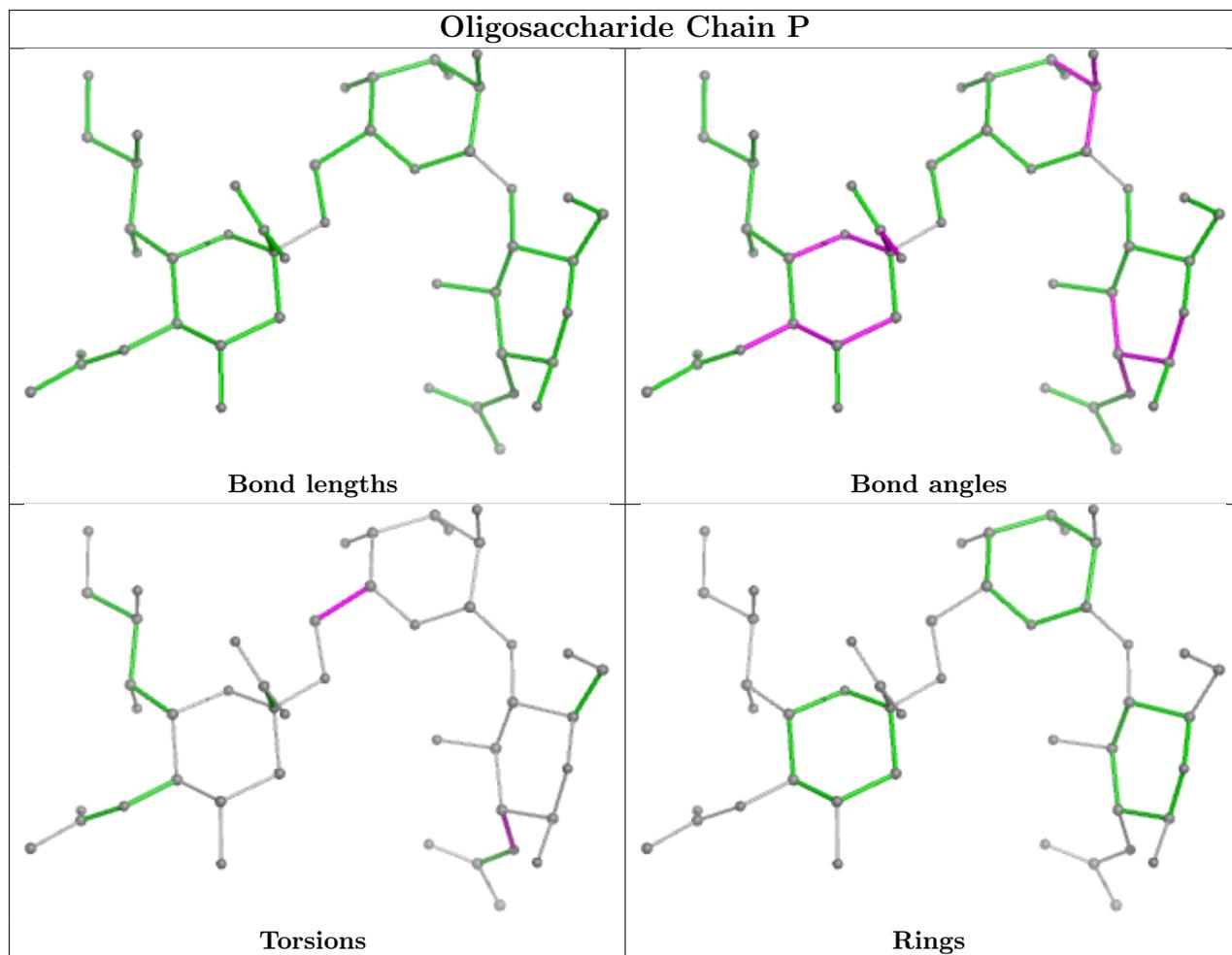
There are no ring outliers.

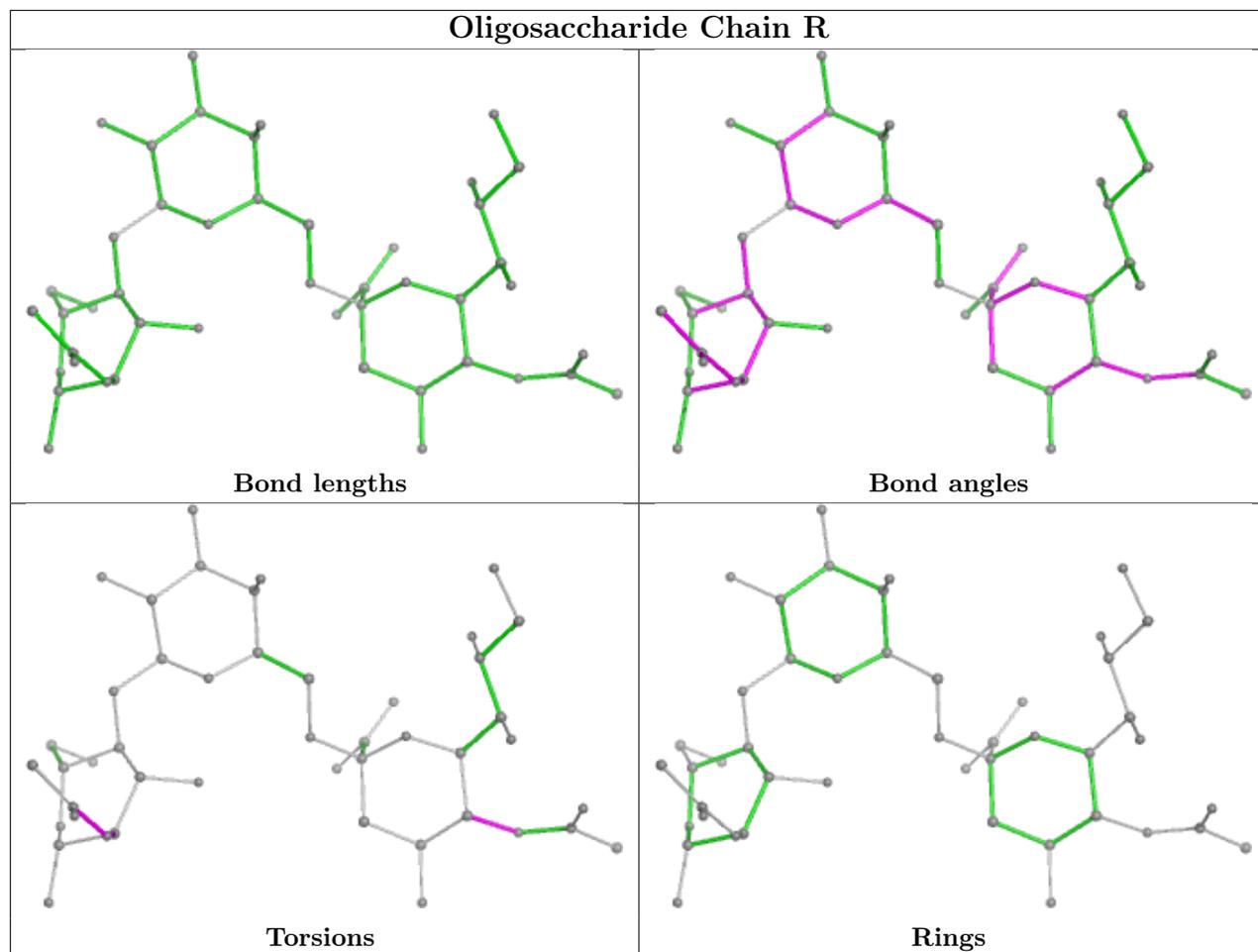
5 monomers are involved in 7 short contacts:

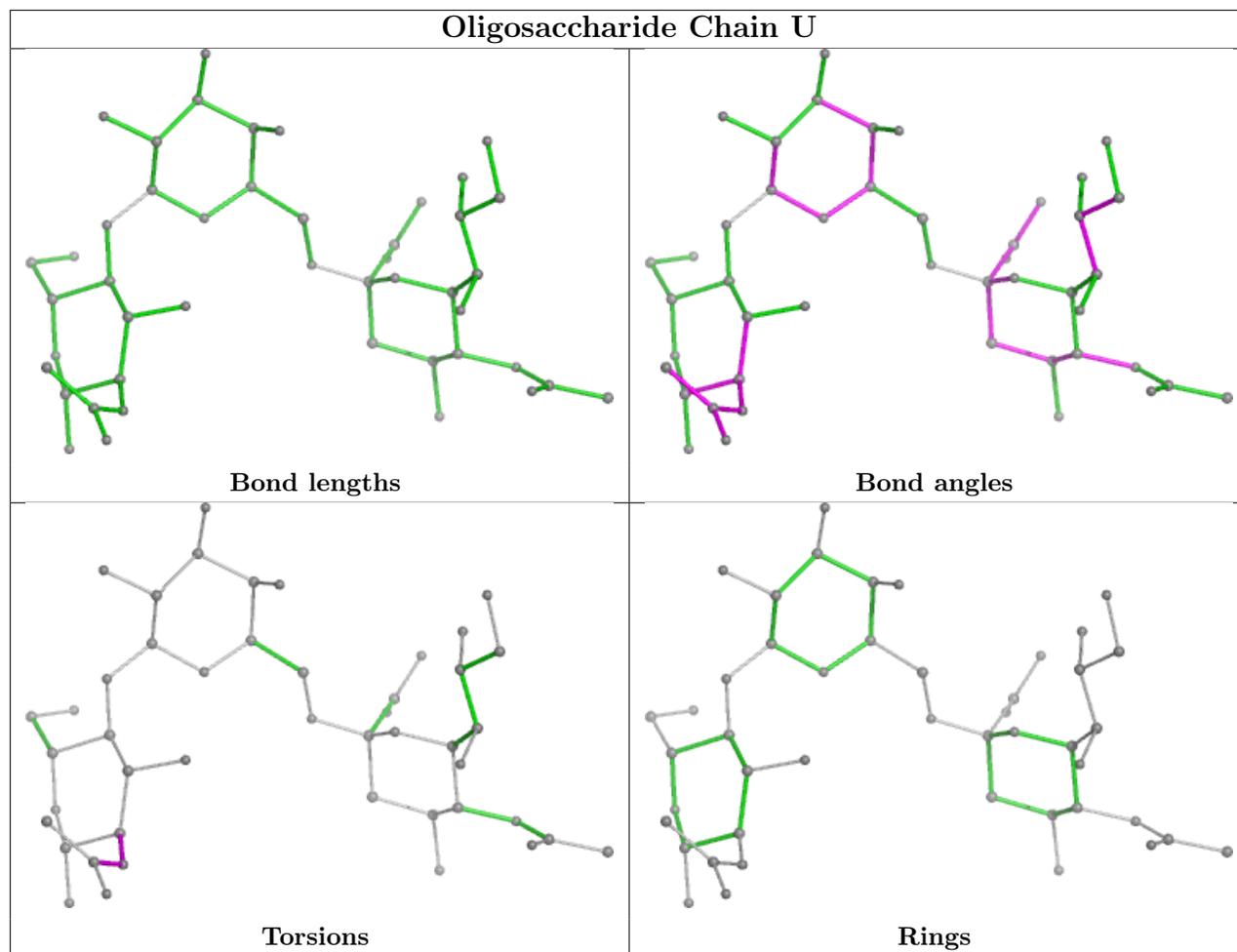
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	M	1	NAG	1	0
3	R	1	NAG	2	0
3	R	2	GAL	2	0
3	U	1	NAG	1	0
4	N	2	SIA	1	0

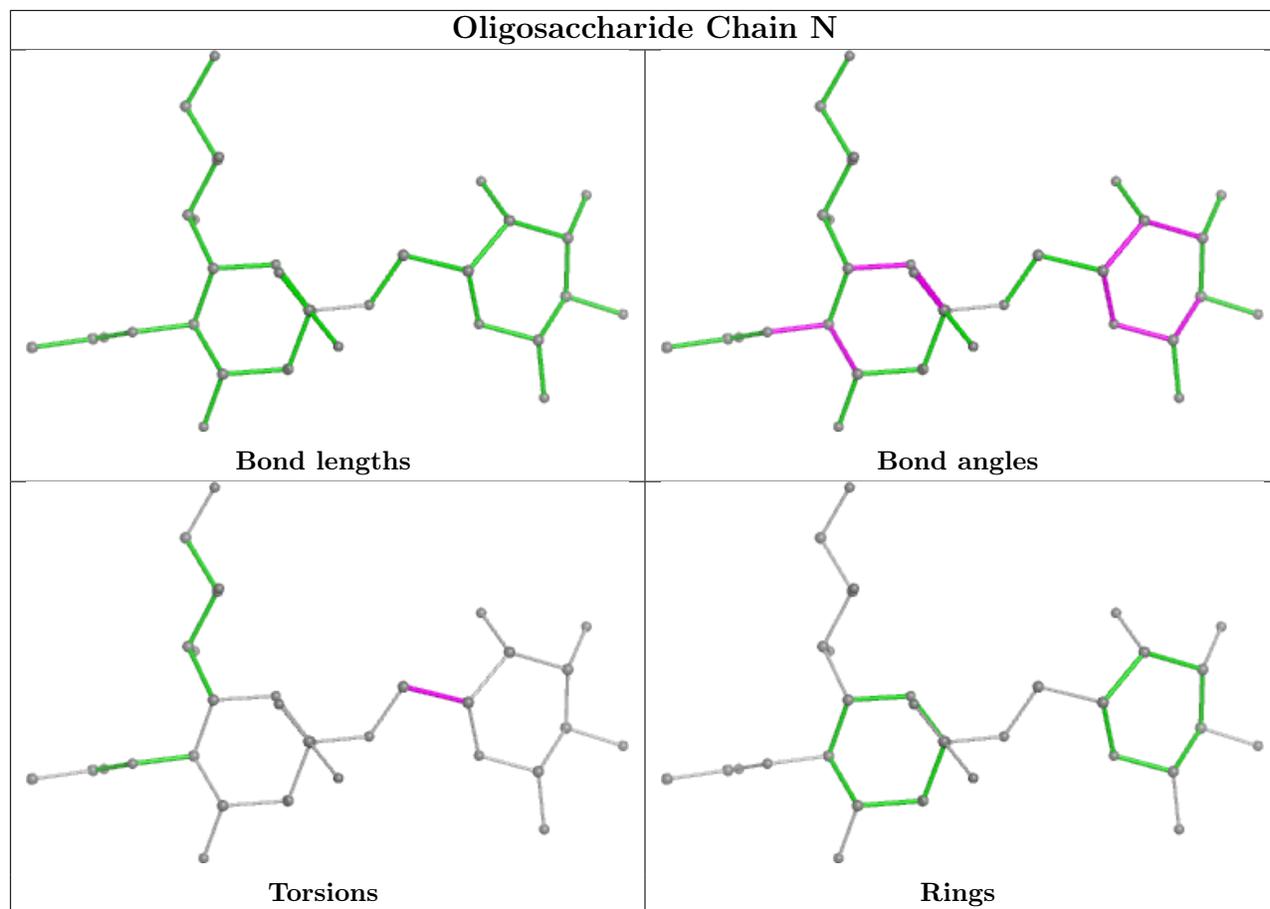
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

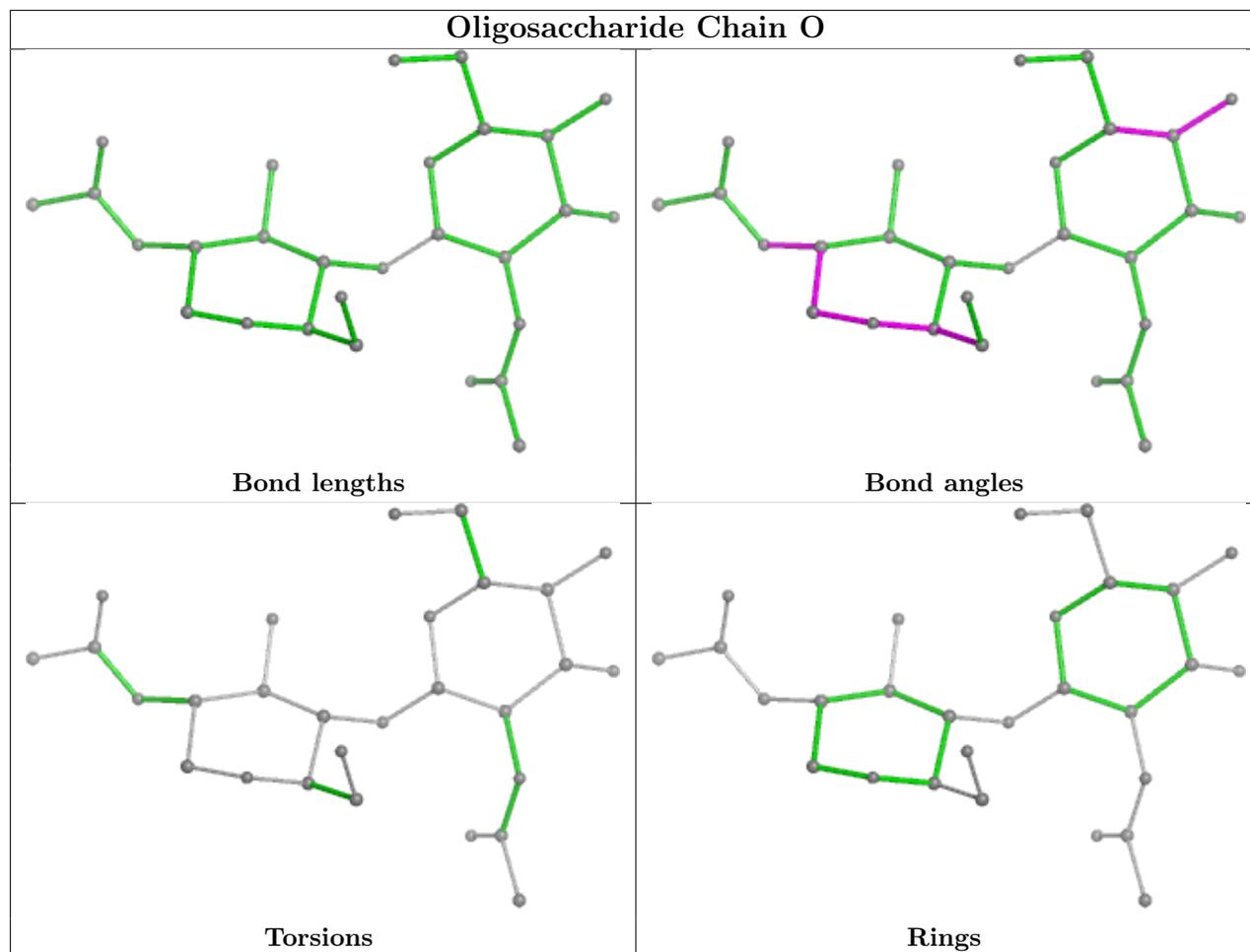


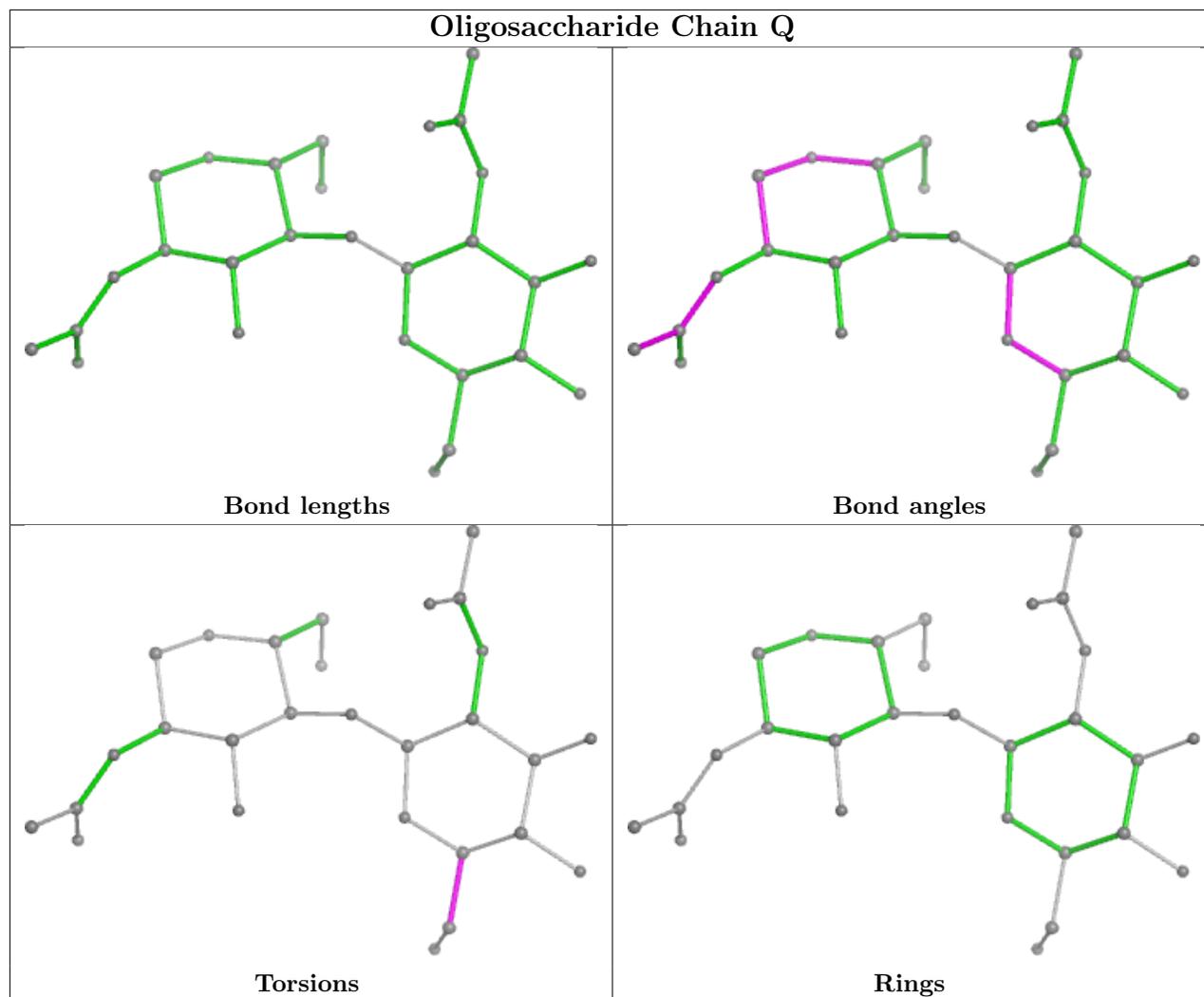


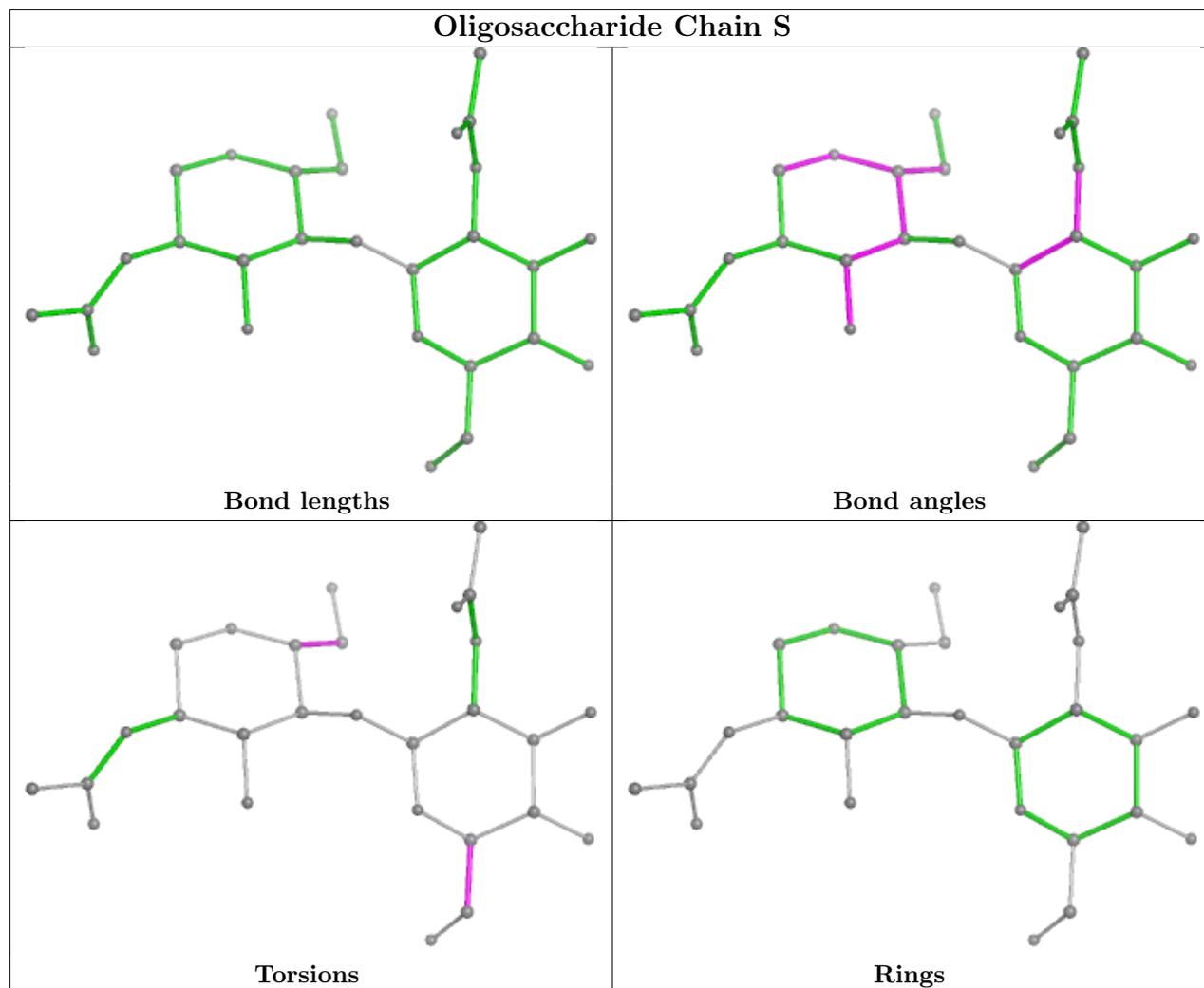


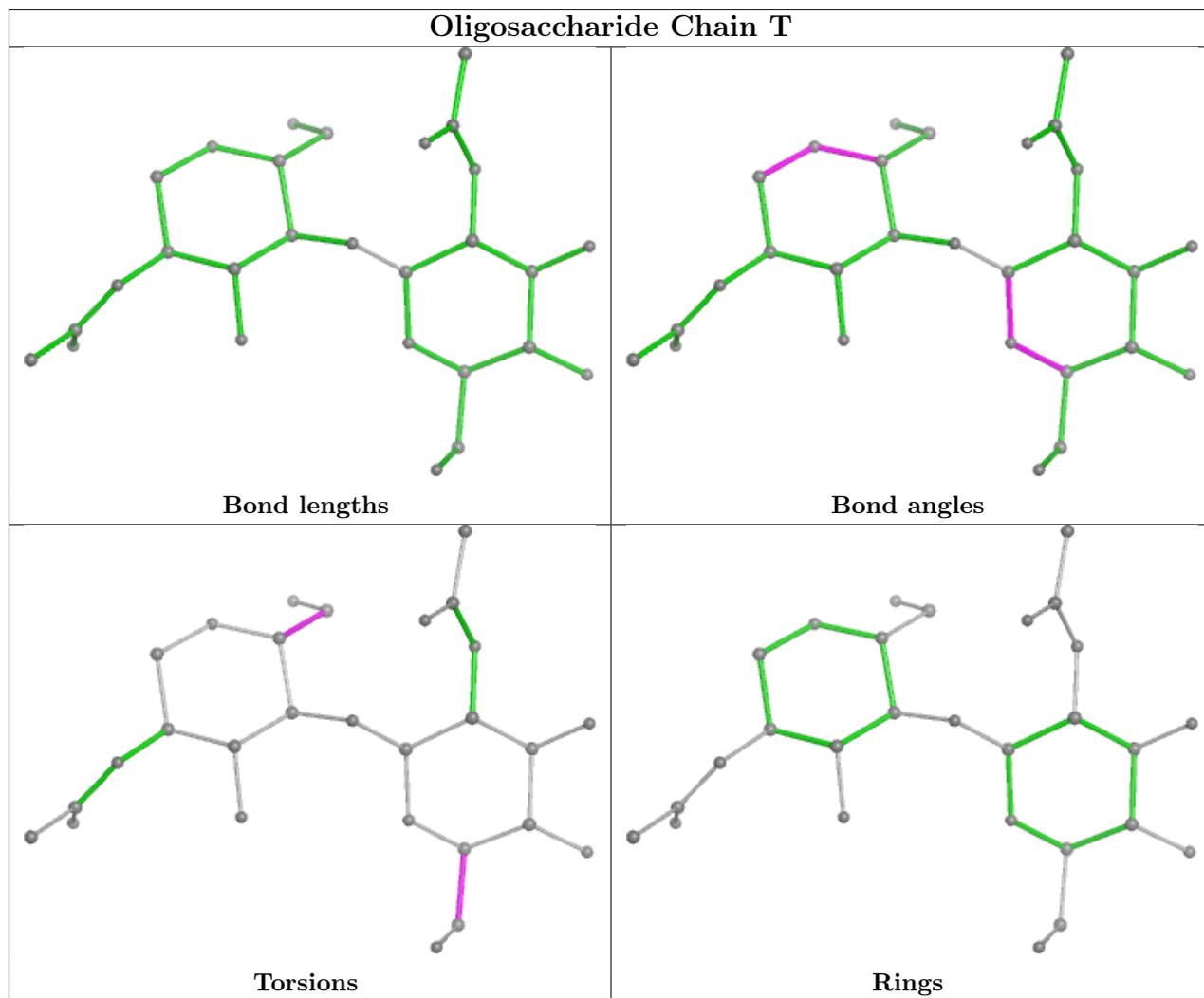


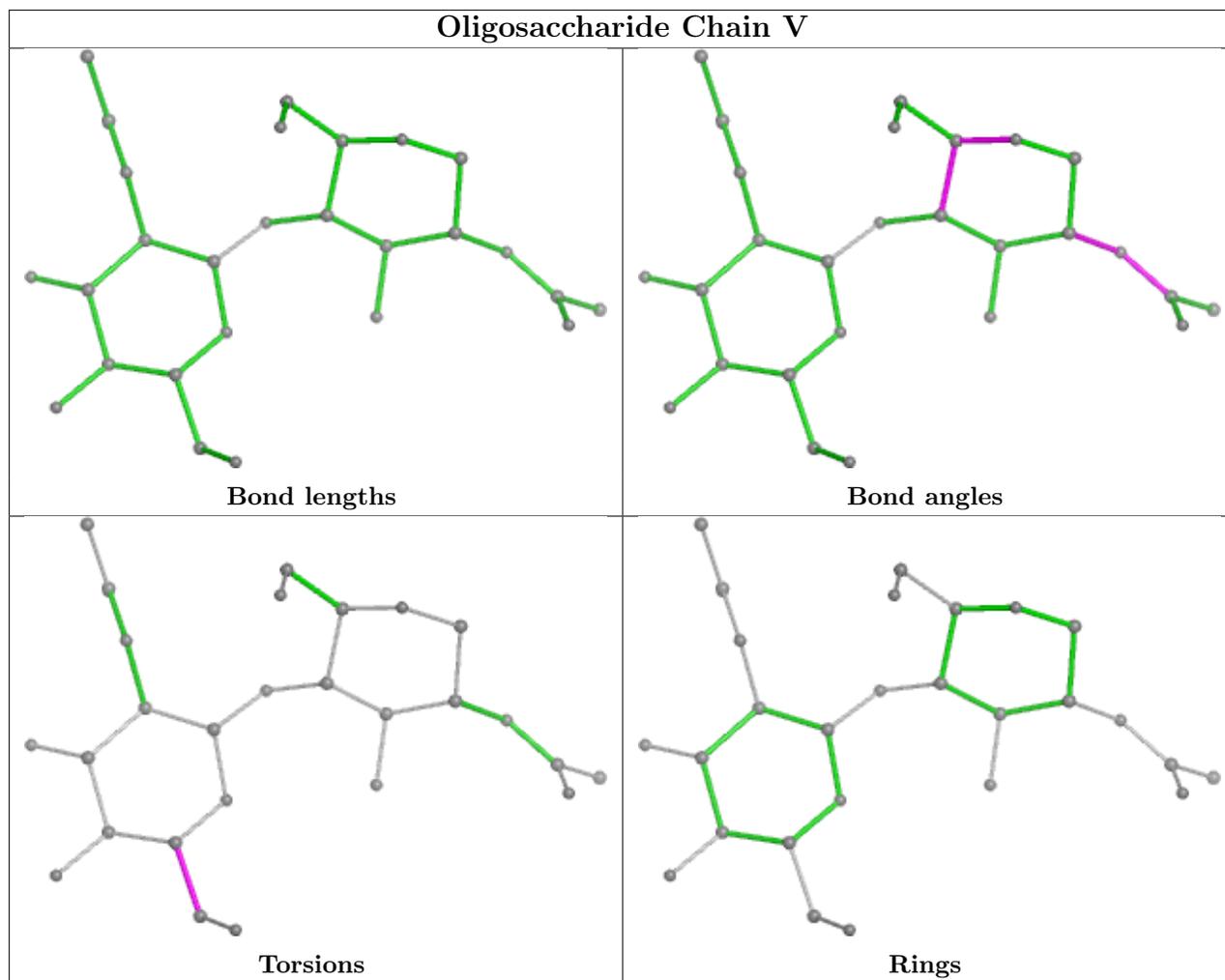












## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 3 are monoatomic - leaving 4 for Mogul analysis.

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	$\# Z  > 2$
7	NAG	L	201	2	14,14,15	0.51	0	17,19,21	1.76	3 (17%)
7	NAG	B	202	2	14,14,15	0.62	1 (7%)	17,19,21	1.35	2 (11%)
8	SIA	I	401	-	21,21,21	0.69	1 (4%)	25,31,31	0.92	2 (8%)
7	NAG	B	201	6,2	14,14,15	0.55	0	17,19,21	1.65	3 (17%)

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
7	NAG	L	201	2	-	2/6/23/26	0/1/1/1
7	NAG	B	202	2	-	2/6/23/26	0/1/1/1
8	SIA	I	401	-	-	6/20/38/38	0/1/1/1
7	NAG	B	201	6,2	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	I	401	SIA	O1B-C1	-2.36	1.21	1.30
7	B	202	NAG	C1-C2	2.02	1.55	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	L	201	NAG	C1-C2-N2	5.12	119.24	110.49
7	B	201	NAG	C1-O5-C5	4.29	118.01	112.19
7	B	201	NAG	O5-C1-C2	-3.58	105.63	111.29
7	L	201	NAG	O5-C1-C2	-3.19	106.25	111.29
7	B	202	NAG	O5-C1-C2	2.99	116.00	111.29

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

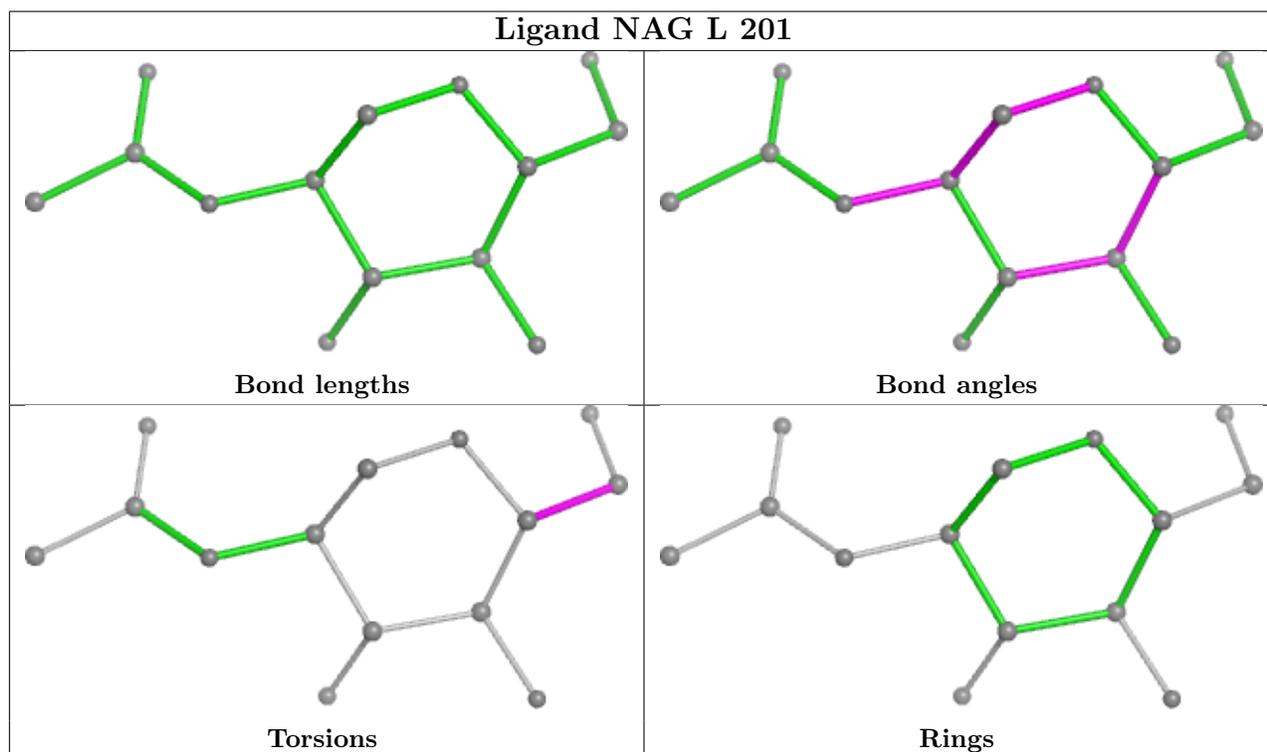
Mol	Chain	Res	Type	Atoms
8	I	401	SIA	O8-C8-C9-O9
8	I	401	SIA	C7-C8-C9-O9
7	B	202	NAG	C4-C5-C6-O6
8	I	401	SIA	O7-C7-C8-C9
8	I	401	SIA	C6-C7-C8-C9

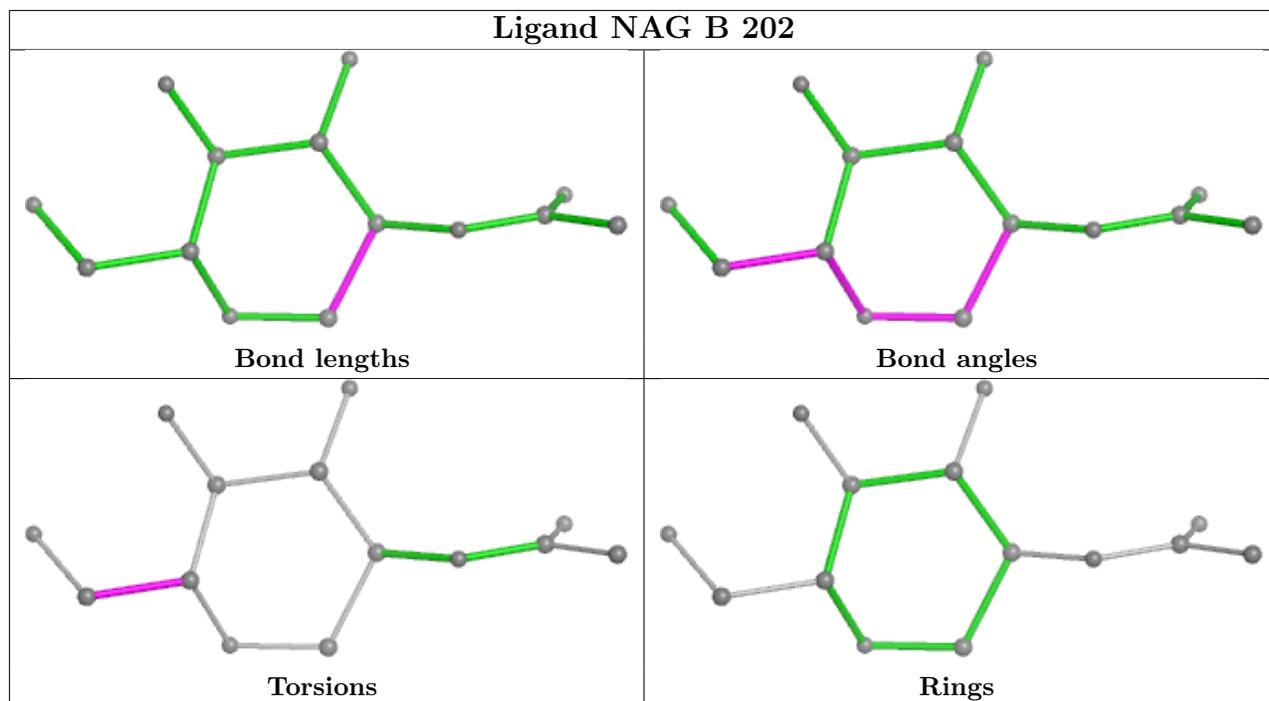
There are no ring outliers.

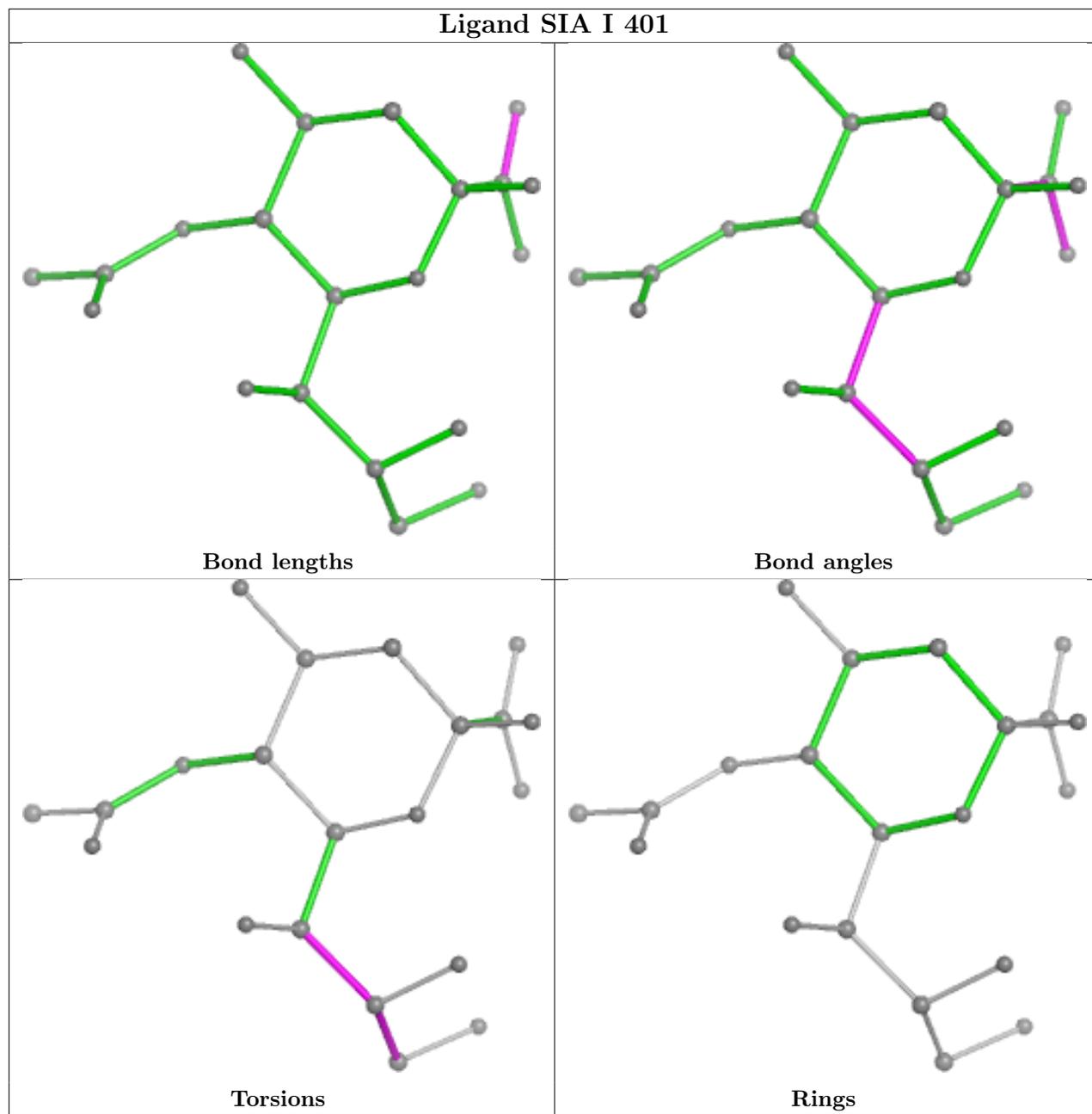
2 monomers are involved in 3 short contacts:

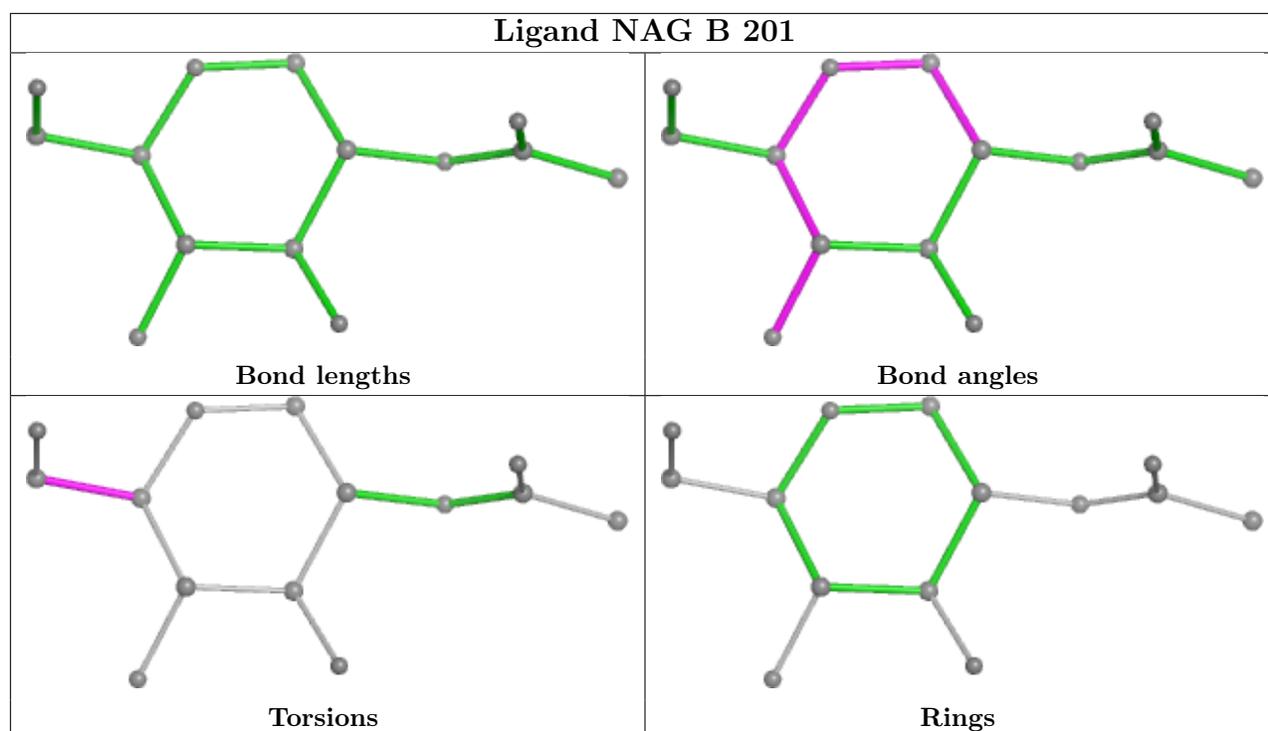
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	L	201	NAG	1	0
8	I	401	SIA	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.









## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [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	320/325 (98%)	-0.12	1 (0%) 90 88	48, 76, 108, 140	1 (0%)
1	C	320/325 (98%)	0.13	5 (1%) 70 65	55, 88, 120, 137	1 (0%)
1	E	320/325 (98%)	0.19	5 (1%) 70 65	58, 98, 129, 146	1 (0%)
1	G	320/325 (98%)	0.87	45 (14%) 7 6	53, 120, 171, 196	0
1	I	320/325 (98%)	0.30	12 (3%) 44 38	62, 108, 155, 183	1 (0%)
1	K	318/325 (97%)	-0.15	2 (0%) 85 83	55, 76, 97, 116	0
2	B	172/177 (97%)	-0.22	0 100 100	58, 81, 110, 122	0
2	D	172/177 (97%)	-0.26	2 (1%) 76 72	52, 75, 99, 119	0
2	F	172/177 (97%)	-0.07	1 (0%) 85 83	55, 84, 112, 131	0
2	H	172/177 (97%)	-0.17	0 100 100	54, 79, 111, 139	0
2	J	172/177 (97%)	0.00	3 (1%) 69 64	61, 85, 115, 132	0
2	L	172/177 (97%)	-0.27	0 100 100	56, 75, 103, 116	0
All	All	2950/3012 (97%)	0.08	76 (2%) 57 51	48, 86, 140, 196	4 (0%)

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	210	VAL	5.2
1	G	211	GLY	5.1
1	G	212	ALA	4.9
1	G	155	PRO	4.8
1	G	188	TYR	4.8

### 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](#)

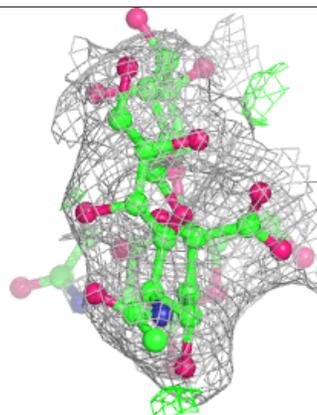
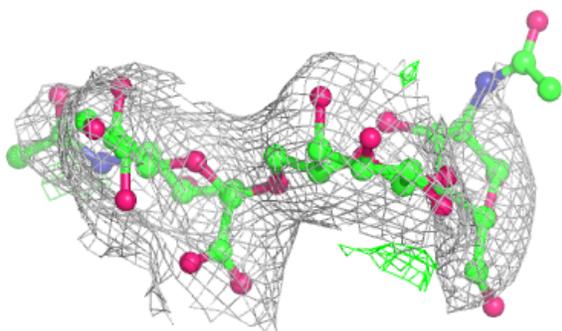
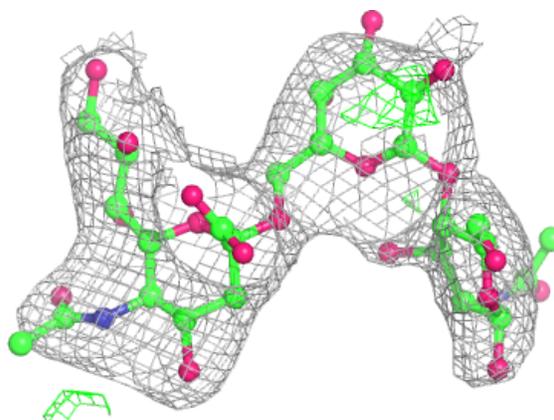
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	GAL	N	1	12/12	0.54	0.13	124,140,153,156	0
3	NAG	R	1	15/15	0.56	0.19	130,142,153,161	15
3	GAL	R	2	11/12	0.64	0.09	146,152,158,160	0
3	NAG	M	1	15/15	0.66	0.12	119,133,147,149	15
5	NAG	Q	1	14/15	0.73	0.11	121,135,141,143	0
5	NAG	V	2	14/15	0.74	0.10	122,131,140,148	0
5	NAG	S	2	14/15	0.75	0.11	117,126,134,137	0
3	NAG	U	1	15/15	0.76	0.13	108,124,137,138	15
3	SIA	R	3	20/21	0.77	0.14	87,123,146,147	0
5	NAG	O	2	14/15	0.79	0.12	107,113,129,130	0
3	NAG	P	1	15/15	0.79	0.11	104,122,131,141	15
5	NAG	Q	2	14/15	0.80	0.10	118,147,152,155	0
3	GAL	M	2	11/12	0.81	0.10	110,124,133,139	0
3	GAL	U	2	11/12	0.81	0.11	101,124,129,130	0
3	GAL	P	2	11/12	0.83	0.09	97,116,135,138	0
5	NAG	V	1	14/15	0.85	0.09	99,113,126,133	0
5	NAG	T	1	14/15	0.87	0.08	110,119,128,134	0
5	NAG	T	2	14/15	0.88	0.09	125,142,148,149	0
5	NAG	S	1	14/15	0.90	0.08	79,94,101,107	0
4	SIA	N	2	20/21	0.92	0.11	93,101,127,138	0
3	SIA	P	3	20/21	0.93	0.08	91,102,120,127	0
5	NAG	O	1	14/15	0.94	0.06	75,84,91,98	0
3	SIA	M	3	20/21	0.95	0.08	61,79,94,95	0
3	SIA	U	3	20/21	0.97	0.07	57,70,97,104	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.

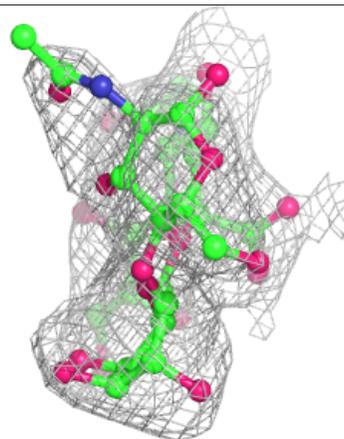
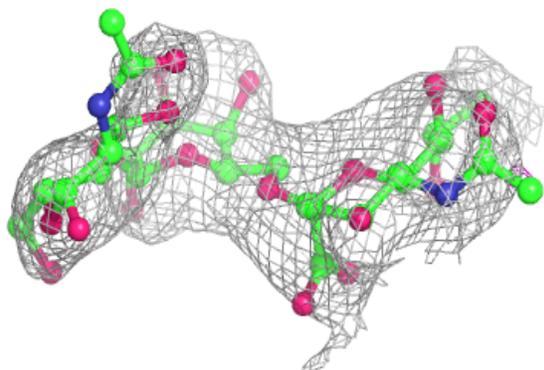
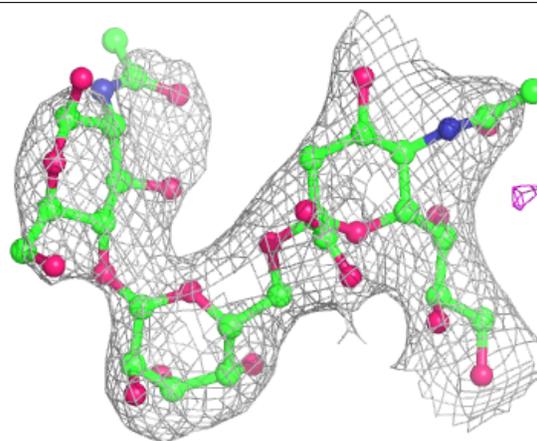
**Electron density around Chain M:**

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

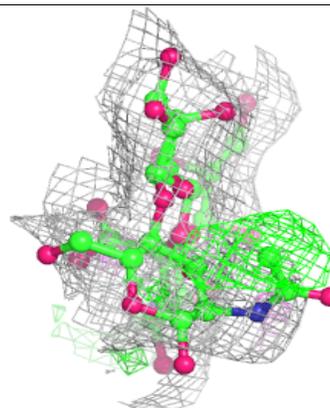
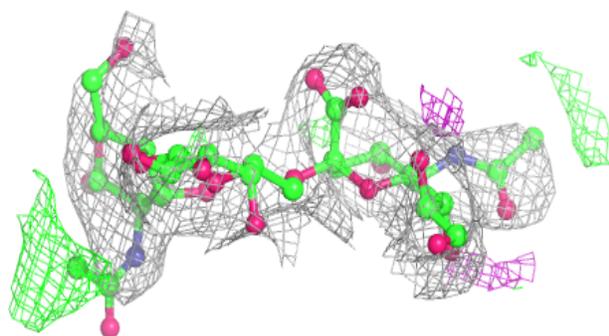
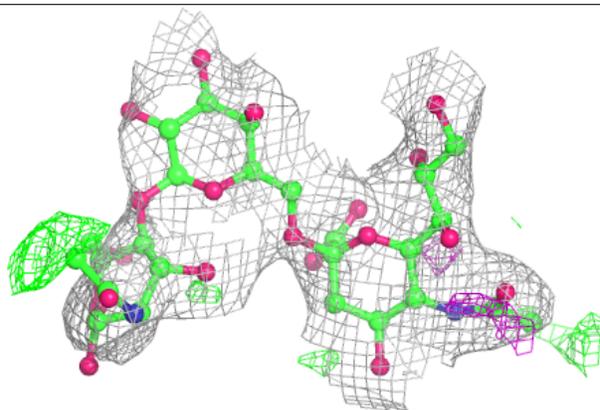


**Electron density around Chain P:**

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

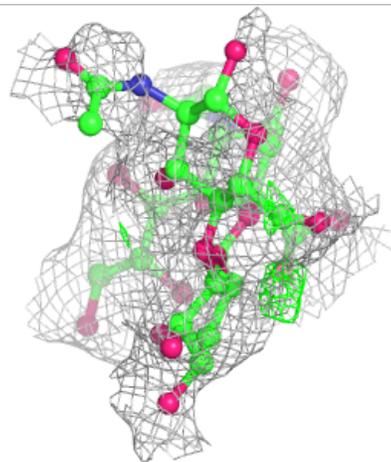
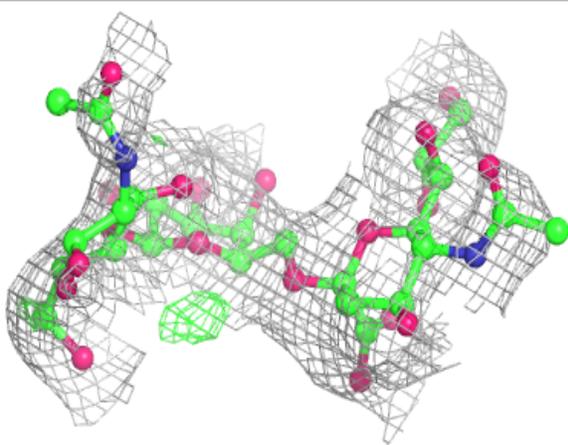
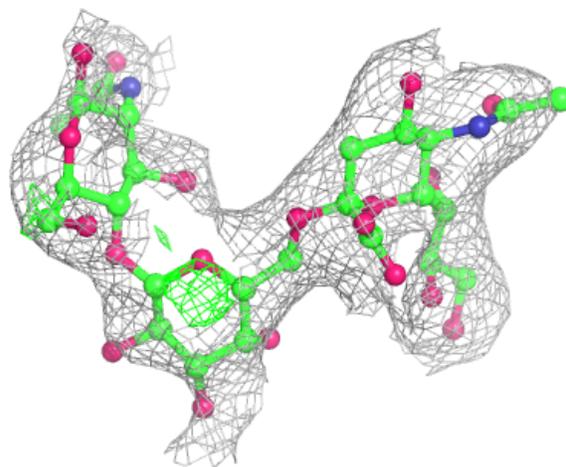
**Electron density around Chain R:**

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



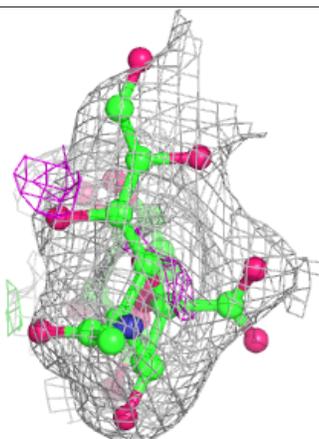
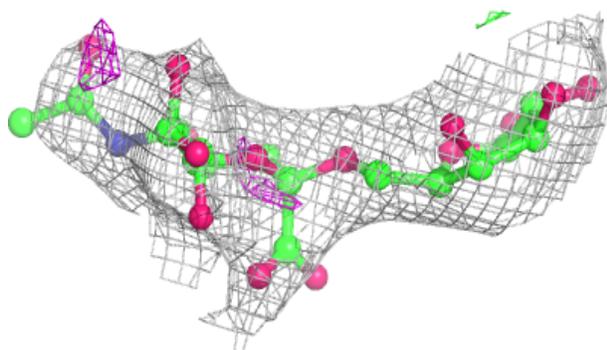
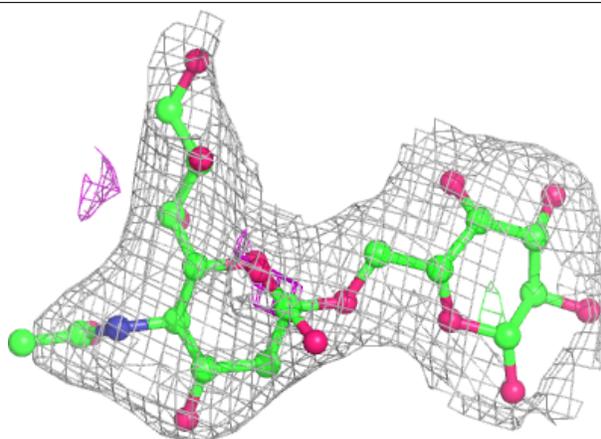
**Electron density around Chain U:**

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

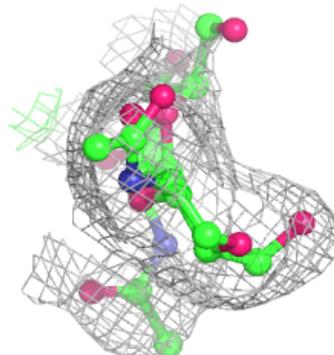
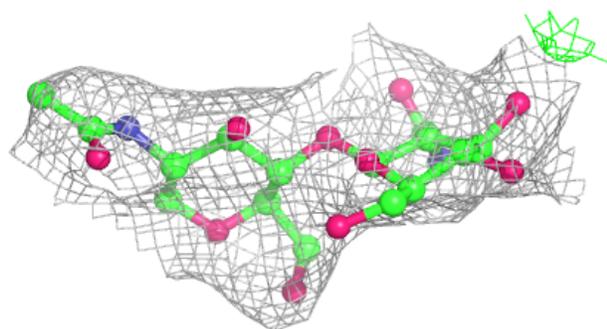
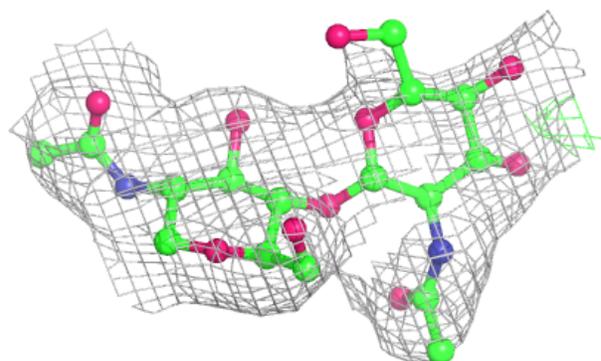


**Electron density around Chain N:**

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

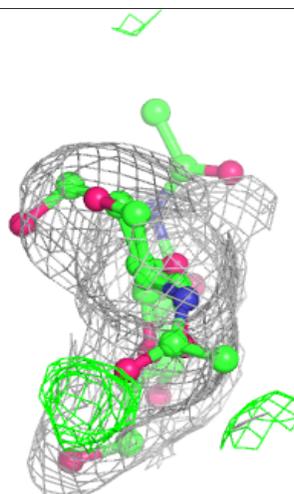
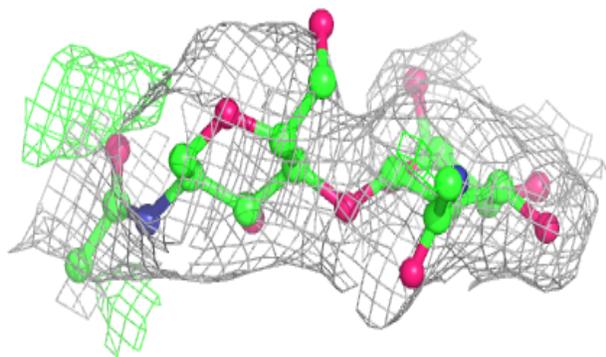
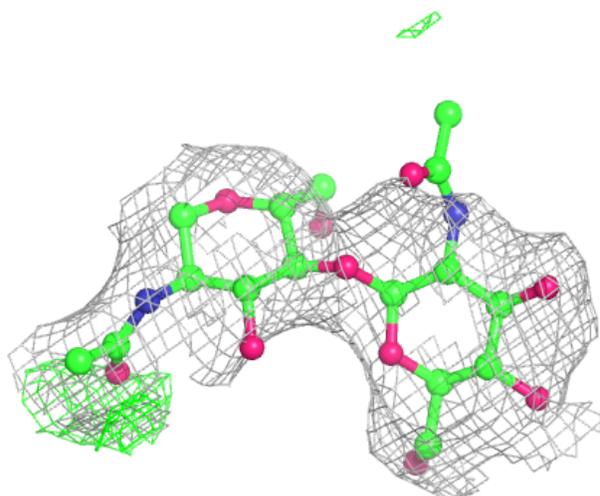
**Electron density around Chain O:**

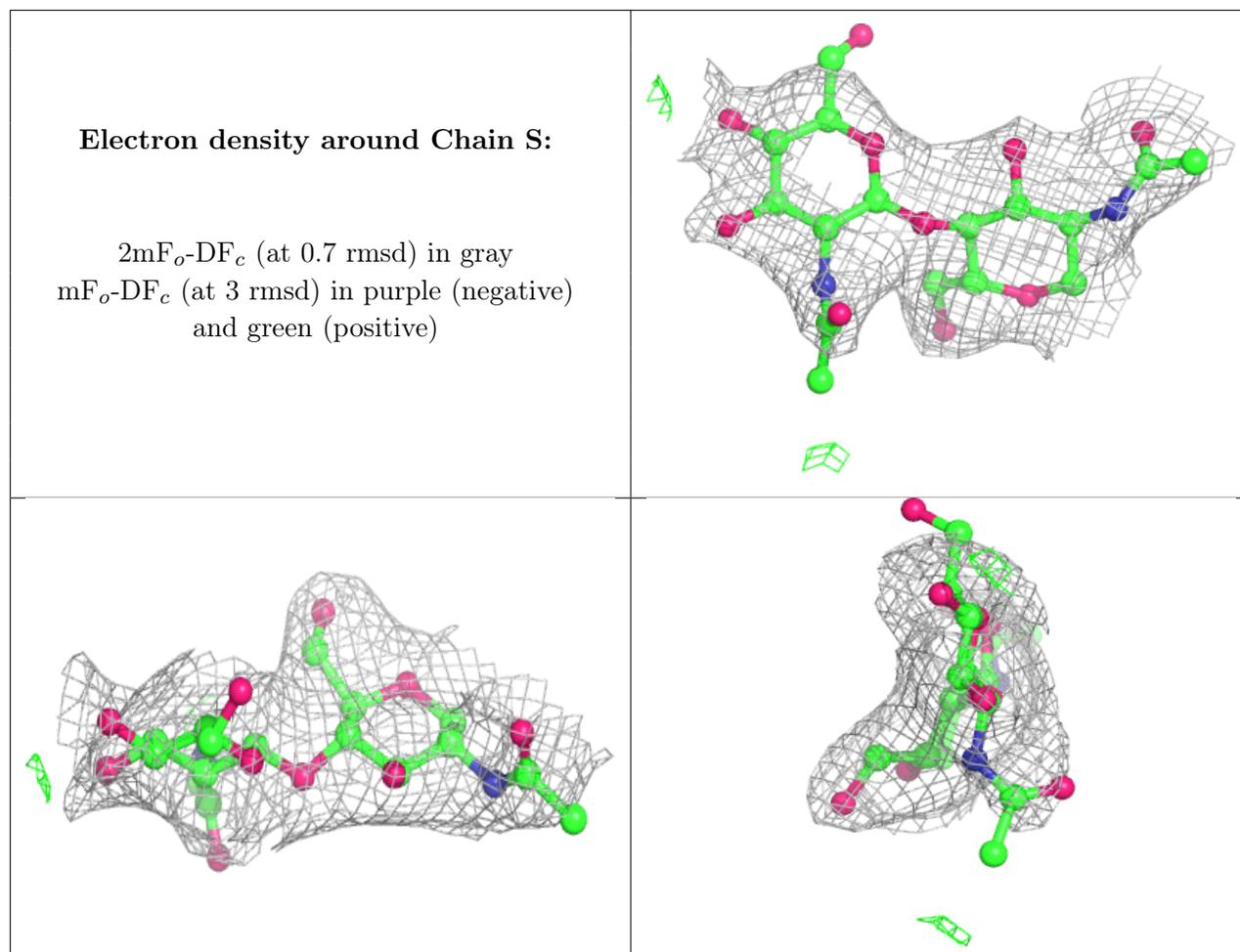
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain Q:**

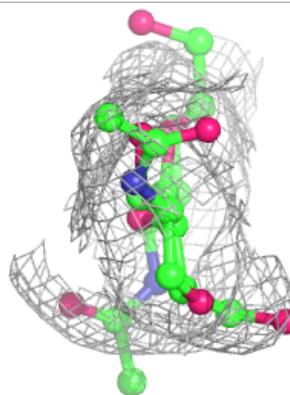
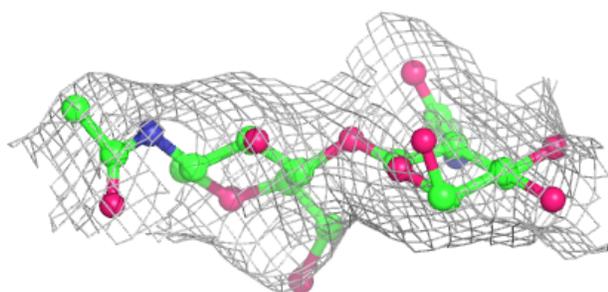
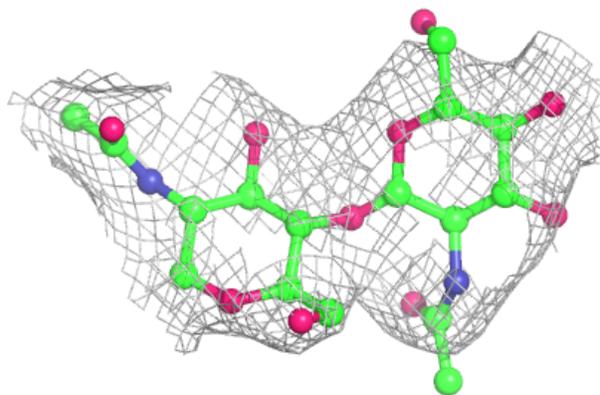
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



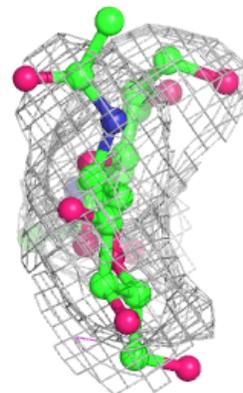
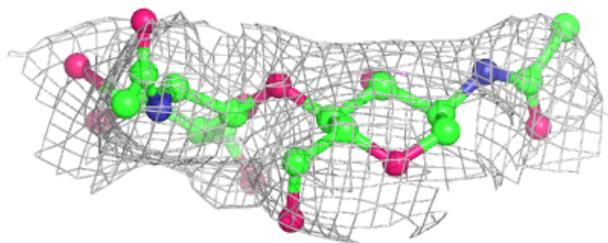
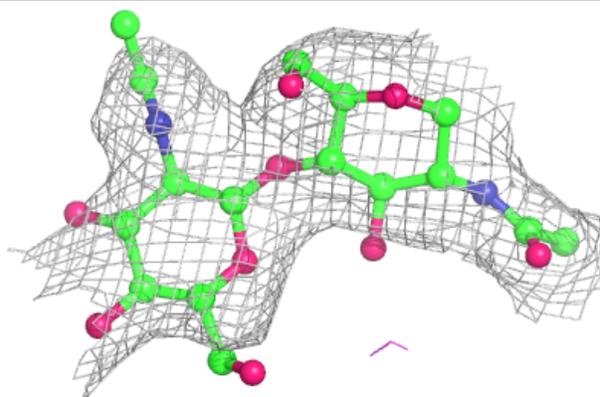


**Electron density around Chain T:**

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

**Electron density around Chain V:**

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

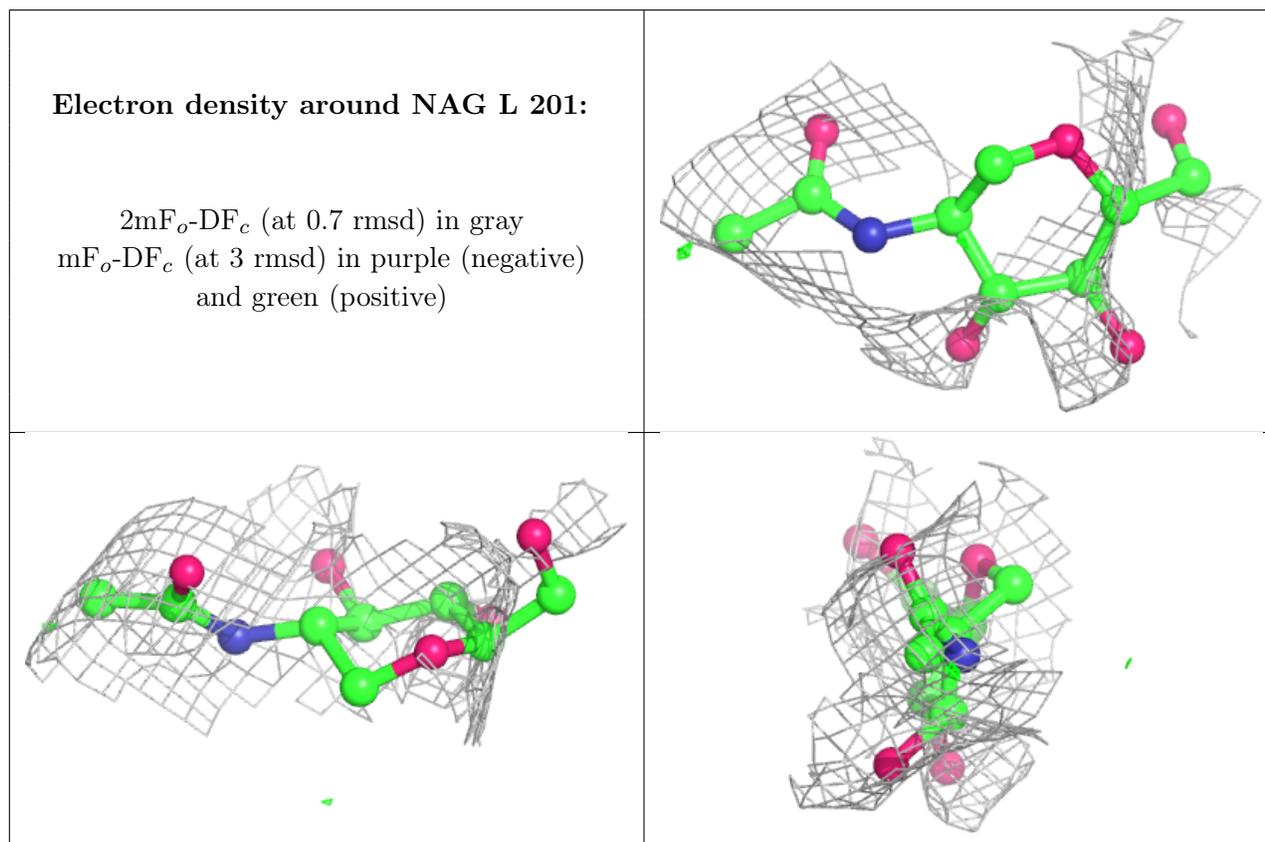


## 6.4 Ligands [i](#)

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.

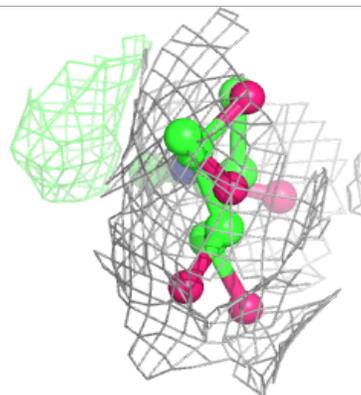
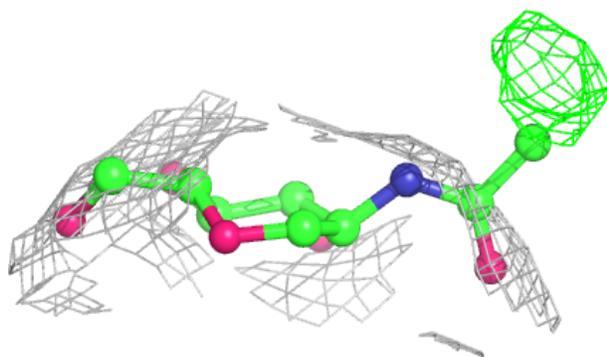
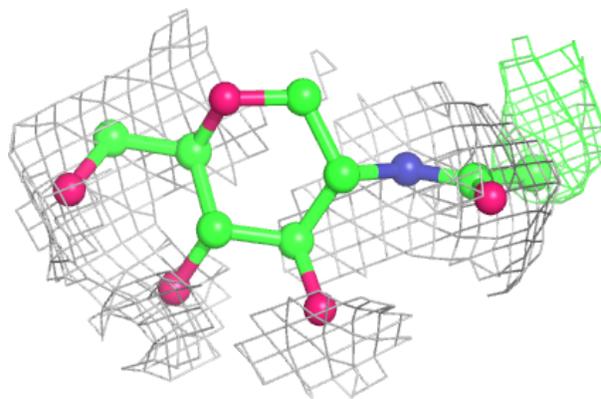
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
7	NAG	L	201	14/15	0.64	0.10	124,161,176,179	0
7	NAG	B	202	14/15	0.67	0.09	150,162,172,175	0
7	NAG	B	201	14/15	0.85	0.09	92,96,101,104	0
8	SIA	I	401	21/21	0.89	0.08	111,137,142,152	0
6	CA	I	402	1/1	0.98	0.08	104,104,104,104	0
6	CA	K	401	1/1	0.98	0.04	67,67,67,67	0
6	CA	A	401	1/1	0.98	0.04	72,72,72,72	0

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

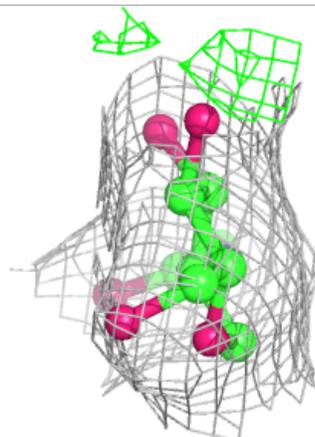
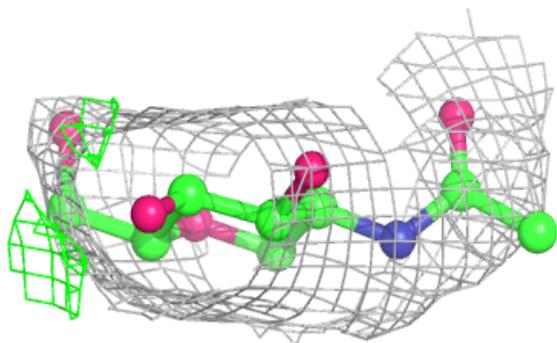
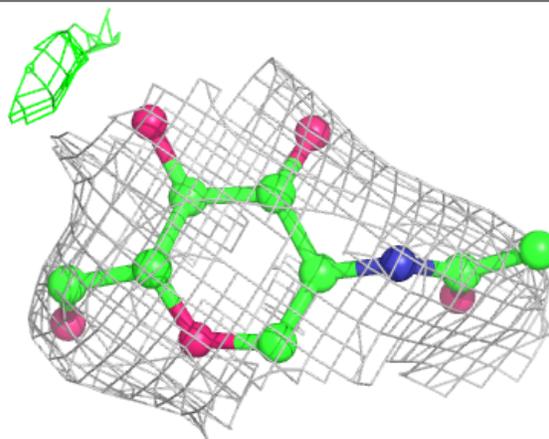


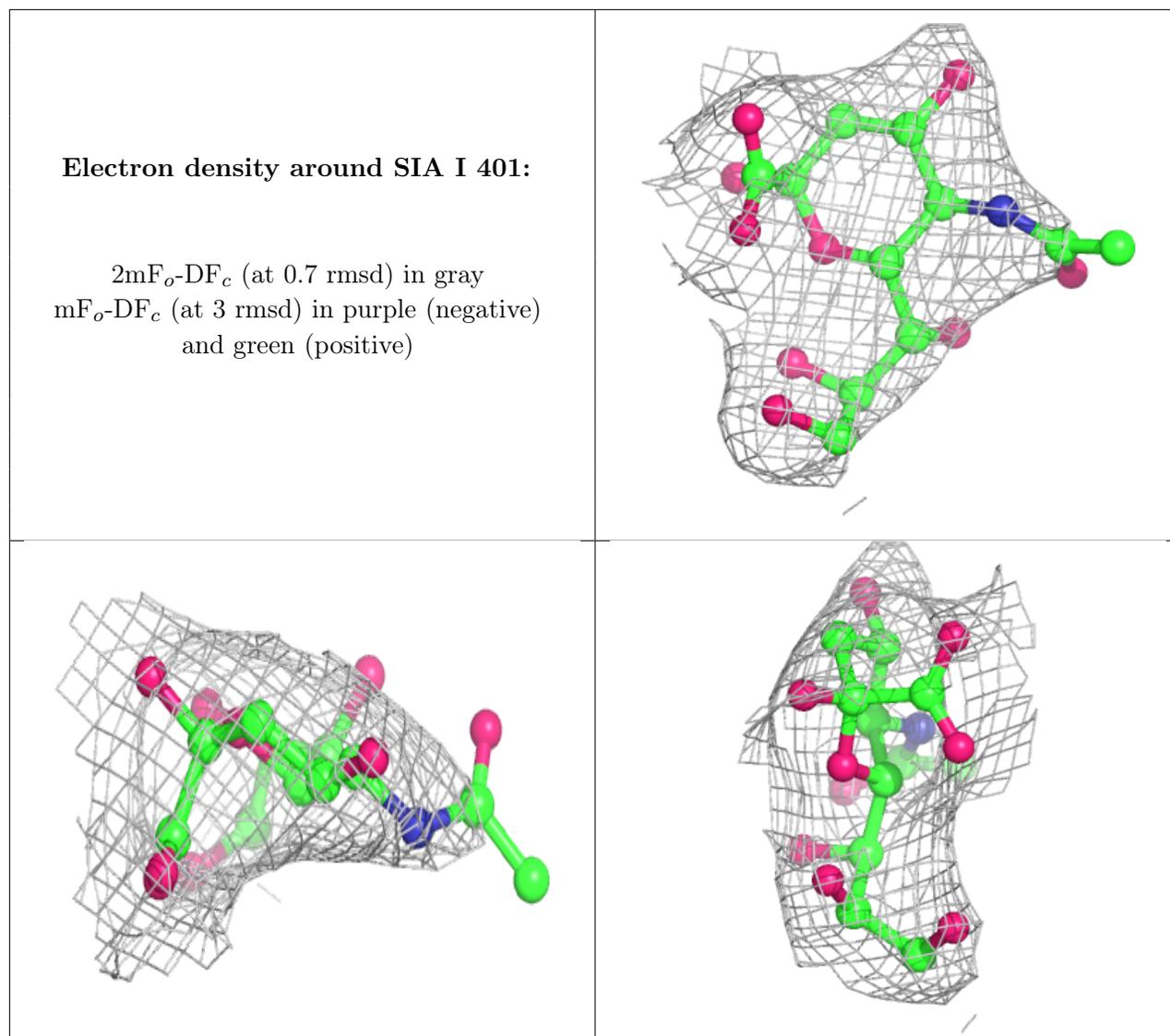
**Electron density around NAG B 202:**

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

**Electron density around NAG B 201:**

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





## 6.5 Other polymers [i](#)

There are no such residues in this entry.