



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 25, 2024 – 12:55 AM EDT

PDB ID : 6B1O  
Title : The structure of DPP4 in complex with Vildagliptin Analog  
Authors : Scapin, G.  
Deposited on : 2017-09-18  
Resolution : 1.91 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.37.1  
buster-report : 1.1.7 (2018)  
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) : 2.37.1

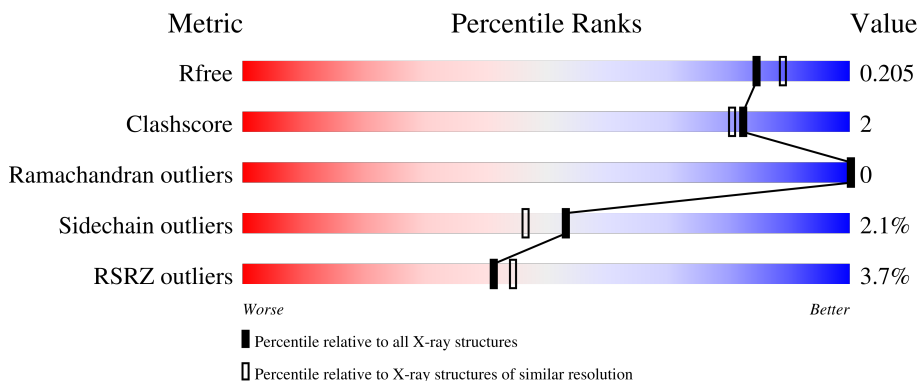
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

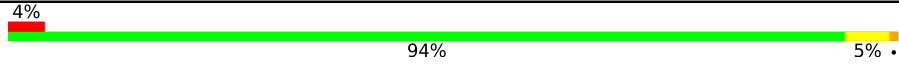
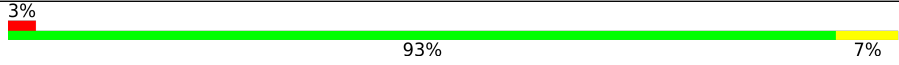
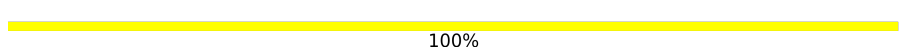
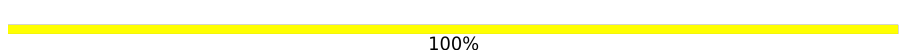
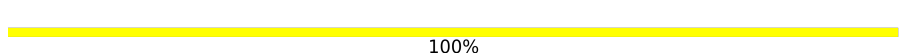
The reported resolution of this entry is 1.91 Å.

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.



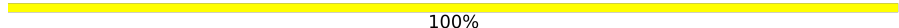
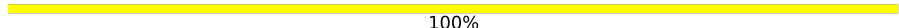
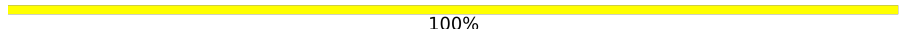


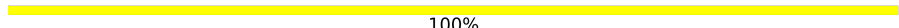
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	728	
1	B	728	
2	C	2	
2	D	2	
2	E	2	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	2	 100%
2	G	2	 100%
2	H	2	 100%
2	I	2	 50% 50%
2	J	2	 100%
2	K	2	 100%

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
2	NAG	C	2	-	-	-	X
2	NAG	E	2	-	-	-	X
2	NAG	F	2	-	-	-	X
2	NAG	G	2	-	-	-	X
2	NAG	H	2	-	-	-	X
2	NAG	K	2	-	-	-	X

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 13880 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 Dipeptidyl peptidase 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	728	6002	3850	988	1138	26	0	8	0
1	B	728	5999	3848	988	1137	26	0	7	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	THR	SER	engineered mutation	UNP P27487
B	39	THR	SER	engineered mutation	UNP P27487

- Molecule 2 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			
2	C	2	28	16	2	10	0	0	0
2	D	2	28	16	2	10	0	0	0
2	E	2	28	16	2	10	0	0	0
2	F	2	28	16	2	10	0	0	0
2	G	2	28	16	2	10	0	0	0
2	H	2	28	16	2	10	0	0	0

*Continued on next page...*

Continued from previous page...

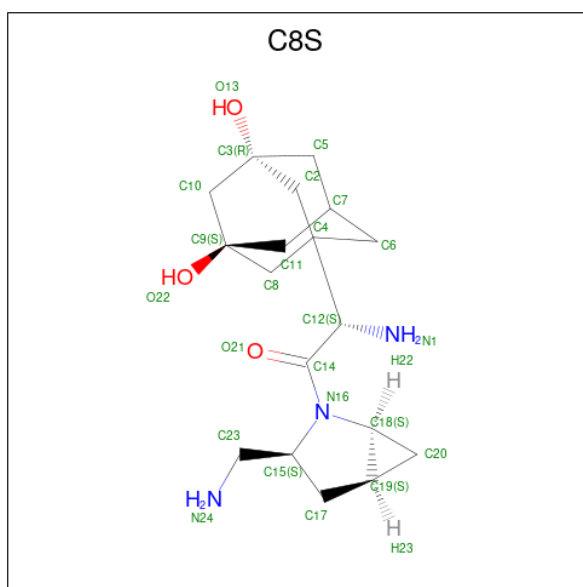
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	I	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	J	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	K	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 3 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		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is (2S)-2-amino-1-[(1S,3S,5S)-3-(aminomethyl)-2-azabicyclo[3.1.0]hexan-2-yl]-2-[(1r,3R,5S,7S)-3,5-dihydroxytricyclo[3.3.1.1 3,7 ]decan-1-yl]ethan-1-one (three-letter code: C8S) (formula: C<sub>18</sub>H<sub>29</sub>N<sub>3</sub>O<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
4	A	1	Total	C	N	O	0	0
			24	18	3	3		
4	B	1	Total	C	N	O	0	0
			24	18	3	3		

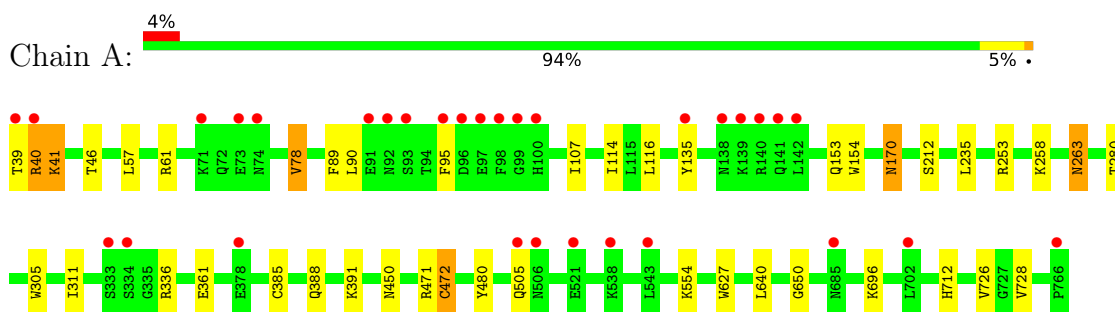
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	744	Total	O	0	0
			744	744		
5	B	793	Total	O	0	0
			793	793		

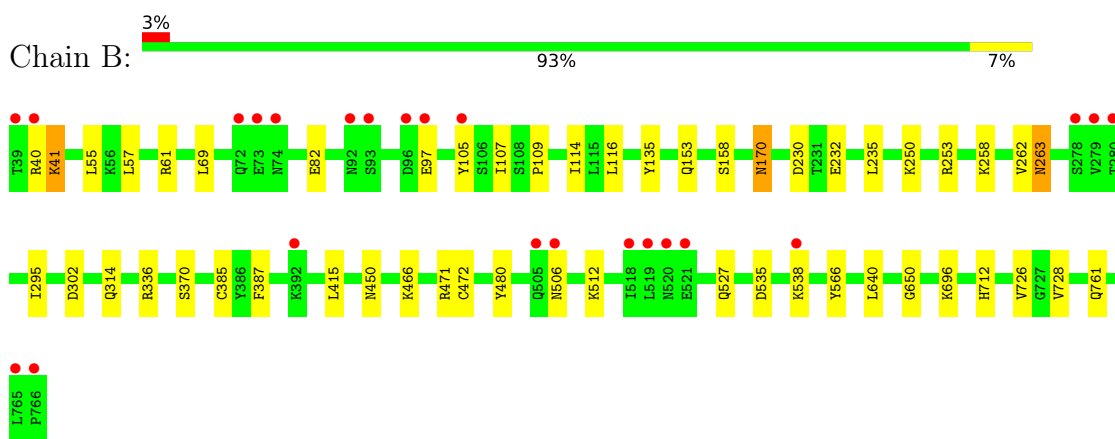
### 3 Residue-property plots [i](#)

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.

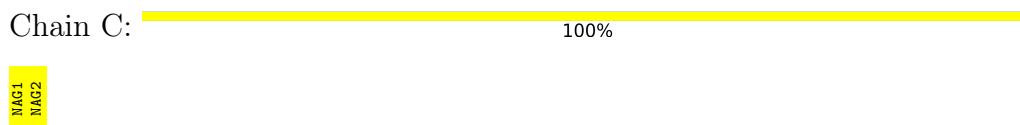
- Molecule 1: Dipeptidyl peptidase 4



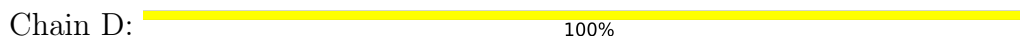
- Molecule 1: Dipeptidyl peptidase 4



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



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



MAG1  
MAG2


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

Chain E:  100%MAG1  
MAG2


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

Chain F:  100%MAG1  
MAG2

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

Chain G:  100%MAG1  
MAG2

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

Chain H:  100%MAG1  
MAG2

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

Chain I:  50% 50%MAG1  
MAG2

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

Chain J:  100%MAG1  
MAG2



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

Chain K:

100%

MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	119.52Å 123.34Å 130.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 1.91 29.93 – 1.91	Depositor EDS
% Data completeness (in resolution range)	100.0 (30.00-1.91) 100.0 (29.93-1.91)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.98 (at 1.91Å)	Xtrriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.172 , 0.202 0.174 , 0.205	Depositor DCC
$R_{free}$ test set	7621 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.0	Xtrriage
Anisotropy	0.042	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 40.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.020 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13880	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.68% 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: C8S, 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	0.50	0/6205	0.57	0/8438
1	B	0.52	0/6199	0.59	0/8430
All	All	0.51	0/12404	0.58	0/16868

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	6002	0	5708	27	0
1	B	5999	0	5703	34	0
2	C	28	0	25	0	0
2	D	28	0	25	0	0
2	E	28	0	25	0	0
2	F	28	0	25	0	0
2	G	28	0	25	0	0
2	H	28	0	25	0	0
2	I	28	0	25	0	0
2	J	28	0	25	1	0
2	K	28	0	25	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	28	0	26	0	0
3	B	14	0	13	0	0
4	A	24	0	0	0	0
4	B	24	0	0	0	0
5	A	744	0	0	6	0
5	B	793	0	0	8	0
All	All	13880	0	11675	59	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 (59) 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:B:450:ASN:OD1	5:B:901:HOH:O	1.77	1.00
1:A:258:LYS:HZ1	1:A:712:HIS:HD2	1.06	0.97
1:A:253:ARG:HH21	1:B:253:ARG:HH21	1.20	0.87
1:A:472:CYS:SG	5:A:935:HOH:O	2.35	0.83
1:B:538:LYS:HE3	5:B:1361:HOH:O	1.83	0.77
1:A:388:GLN:HB2	1:A:391:LYS:HG2	1.69	0.74
1:A:258:LYS:HZ1	1:A:712:HIS:CD2	1.99	0.70
1:B:105:TYR:HB2	1:B:114:ILE:HD11	1.73	0.69
1:B:153:GLN:HE22	1:B:170:ASN:ND2	1.94	0.66
1:B:41:LYS:HD2	1:B:41:LYS:H	1.60	0.66
1:B:471[A]:ARG:HG3	1:B:480:TYR:CE2	2.30	0.66
1:A:153:GLN:HE22	1:A:170:ASN:ND2	1.97	0.63
1:A:78:VAL:HG22	1:A:89:PHE:HB2	1.82	0.62
1:B:258:LYS:NZ	1:B:712:HIS:HD2	1.97	0.62
1:A:90:LEU:HD21	1:A:95:PHE:HE2	1.66	0.60
1:B:696:LYS:HG3	1:B:728:VAL:HG22	1.83	0.60
1:A:726:VAL:HG23	1:A:728:VAL:HG23	1.84	0.60
1:A:471[A]:ARG:HG3	1:A:480:TYR:CE2	2.38	0.59
1:A:253:ARG:NH2	1:B:253:ARG:HH21	1.98	0.58
1:B:761:GLN:NE2	5:B:904:HOH:O	2.37	0.58
1:A:61:ARG:HH22	1:A:107:ILE:H	1.55	0.55
1:A:41:LYS:H	1:A:41:LYS:HD2	1.71	0.55
1:A:696:LYS:HG2	1:A:728:VAL:HG22	1.88	0.55
1:A:46:THR:HG23	5:A:1530:HOH:O	2.08	0.54
1:B:69:LEU:HD13	1:B:107:ILE:HD12	1.90	0.54
1:B:466:LYS:HE3	5:B:1607:HOH:O	2.08	0.53
2:J:1:NAG:H62	2:J:2:NAG:O5	2.09	0.52

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:170:ASN:HD22	1:B:170:ASN:N	2.08	0.52
1:B:535:ASP:HB3	1:B:538:LYS:HD2	1.91	0.51
1:A:640:LEU:HD11	1:A:650:GLY:HA3	1.93	0.51
1:A:253:ARG:HH21	1:B:253:ARG:NH2	1.99	0.49
1:B:370:SER:HB2	1:B:387:PHE:O	2.12	0.49
1:A:258:LYS:NZ	1:A:712:HIS:HD2	1.91	0.49
1:B:258:LYS:HZ3	1:B:712:HIS:HD2	1.61	0.48
1:B:471[B]:ARG:NH2	5:B:916:HOH:O	2.47	0.47
1:A:554:LYS:HG2	5:A:905:HOH:O	2.15	0.47
1:B:640:LEU:HD11	1:B:650:GLY:HA3	1.96	0.46
1:B:512:LYS:HE3	1:B:527:GLN:OE1	2.16	0.45
1:B:302:ASP:HB3	1:B:314:GLN:HB2	1.98	0.45
1:A:40:ARG:NH1	1:A:505:GLN:O	2.50	0.44
1:B:114:ILE:HG23	1:B:135:TYR:HB3	1.98	0.44
1:A:154:TRP:CE2	1:A:212[A]:SER:HB3	2.52	0.44
1:A:471[B]:ARG:NH2	5:A:922:HOH:O	2.51	0.43
1:B:82:GLU:HG2	5:B:1588:HOH:O	2.18	0.43
1:B:258:LYS:HZ3	1:B:712:HIS:CD2	2.36	0.43
1:B:263:ASN:HD22	1:B:263:ASN:HA	1.61	0.42
1:A:263:ASN:HD22	1:A:263:ASN:HA	1.68	0.42
1:B:336:ARG:HD3	5:B:1243:HOH:O	2.19	0.42
1:A:450:ASN:HB2	5:A:1064:HOH:O	2.19	0.42
1:A:305:TRP:CE2	1:A:311:ILE:HD12	2.55	0.42
1:B:295:ILE:H	1:B:295:ILE:HG13	1.71	0.42
1:B:109:PRO:HG2	1:B:158:SER:O	2.20	0.42
1:A:114:ILE:HG23	1:A:135:TYR:HB3	2.01	0.41
1:B:40:ARG:HB2	1:B:506:ASN:O	2.20	0.41
1:B:250:LYS:NZ	5:B:935:HOH:O	2.53	0.41
1:A:253:ARG:HD3	5:A:1554:HOH:O	2.19	0.41
1:B:415:LEU:C	1:B:415:LEU:HD23	2.41	0.41
1:B:726:VAL:HG23	1:B:728:VAL:HG23	2.02	0.41
1:B:232:GLU:HB2	1:B:262:VAL:HG11	2.02	0.40

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	734/728 (101%)	712 (97%)	22 (3%)	0	100	100
1	B	733/728 (101%)	710 (97%)	23 (3%)	0	100	100
All	All	1467/1456 (101%)	1422 (97%)	45 (3%)	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	661/653 (101%)	645 (98%)	16 (2%)	49	41
1	B	660/653 (101%)	647 (98%)	13 (2%)	55	49
All	All	1321/1306 (101%)	1292 (98%)	29 (2%)	53	45

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	39	THR
1	A	40	ARG
1	A	41	LYS
1	A	57	LEU
1	A	78	VAL
1	A	116	LEU
1	A	170	ASN
1	A	235	LEU
1	A	263	ASN
1	A	280	THR
1	A	336	ARG
1	A	361[A]	GLU
1	A	361[B]	GLU
1	A	385	CYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	472	CYS
1	A	627	TRP
1	B	41	LYS
1	B	55	LEU
1	B	57	LEU
1	B	61	ARG
1	B	97	GLU
1	B	116	LEU
1	B	170	ASN
1	B	230	ASP
1	B	235	LEU
1	B	263	ASN
1	B	385	CYS
1	B	472	CYS
1	B	566	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	75	ASN
1	A	169	ASN
1	A	170	ASN
1	A	263	ASN
1	A	572	ASN
1	A	694	ASN
1	A	712	HIS
1	B	141	GLN
1	B	169	ASN
1	B	170	ASN
1	B	263	ASN
1	B	430	ASN
1	B	435	GLN
1	B	572	ASN
1	B	685	ASN
1	B	694	ASN
1	B	712	HIS
1	B	761	GLN

### 5.3.3 RNA

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](#)

18 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
2	NAG	C	1	1,2	14,14,15	0.50	0	17,19,21	1.06	2 (11%)
2	NAG	C	2	2	14,14,15	0.77	0	17,19,21	1.48	3 (17%)
2	NAG	D	1	1,2	14,14,15	0.63	0	17,19,21	1.02	1 (5%)
2	NAG	D	2	2	14,14,15	0.50	0	17,19,21	0.97	2 (11%)
2	NAG	E	1	1,2	14,14,15	0.55	0	17,19,21	1.59	3 (17%)
2	NAG	E	2	2	14,14,15	0.82	0	17,19,21	1.37	3 (17%)
2	NAG	F	1	1,2	14,14,15	0.62	0	17,19,21	1.08	1 (5%)
2	NAG	F	2	2	14,14,15	0.54	0	17,19,21	2.01	3 (17%)
2	NAG	G	1	1,2	14,14,15	0.47	0	17,19,21	1.81	1 (5%)
2	NAG	G	2	2	14,14,15	0.49	0	17,19,21	1.09	1 (5%)
2	NAG	H	1	1,2	14,14,15	0.49	0	17,19,21	1.10	1 (5%)
2	NAG	H	2	2	14,14,15	0.61	0	17,19,21	2.36	5 (29%)
2	NAG	I	1	1,2	14,14,15	0.57	0	17,19,21	0.88	0
2	NAG	I	2	2	14,14,15	0.41	0	17,19,21	2.07	4 (23%)
2	NAG	J	1	1,2	14,14,15	0.42	0	17,19,21	1.76	2 (11%)
2	NAG	J	2	2	14,14,15	0.87	1 (7%)	17,19,21	1.53	2 (11%)
2	NAG	K	1	1,2	14,14,15	0.57	0	17,19,21	1.14	1 (5%)
2	NAG	K	2	2	14,14,15	0.52	0	17,19,21	1.82	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
2	NAG	C	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	C	2	2	-	0/6/23/26	0/1/1/1
2	NAG	D	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	D	2	2	-	2/6/23/26	0/1/1/1
2	NAG	E	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	E	2	2	-	3/6/23/26	0/1/1/1
2	NAG	F	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	F	2	2	-	3/6/23/26	0/1/1/1
2	NAG	G	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	G	2	2	-	2/6/23/26	0/1/1/1
2	NAG	H	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	H	2	2	-	3/6/23/26	0/1/1/1
2	NAG	I	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	I	2	2	-	2/6/23/26	0/1/1/1
2	NAG	J	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	J	2	2	-	1/6/23/26	0/1/1/1
2	NAG	K	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	K	2	2	-	4/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	2	NAG	O5-C1	-2.14	1.40	1.43

All (38) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	1	NAG	C1-O5-C5	6.79	121.39	112.19
2	I	2	NAG	C1-O5-C5	6.73	121.31	112.19
2	H	2	NAG	C2-N2-C7	5.93	131.34	122.90
2	K	2	NAG	C1-O5-C5	5.92	120.21	112.19
2	F	2	NAG	C1-O5-C5	5.51	119.66	112.19
2	H	2	NAG	C1-O5-C5	5.15	119.17	112.19
2	J	1	NAG	O4-C4-C3	-4.83	99.19	110.35
2	E	1	NAG	O4-C4-C3	-4.68	99.52	110.35
2	F	2	NAG	O5-C5-C6	4.63	114.46	107.20
2	J	2	NAG	C2-N2-C7	4.27	128.98	122.90
2	J	1	NAG	C1-O5-C5	4.10	117.74	112.19
2	C	2	NAG	C4-C3-C2	3.91	116.74	111.02
2	H	2	NAG	O5-C5-C6	3.37	112.49	107.20
2	K	1	NAG	O5-C1-C2	-2.92	106.68	111.29

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	2	NAG	C1-C2-N2	2.69	115.08	110.49
2	E	2	NAG	C2-N2-C7	2.68	126.72	122.90
2	G	2	NAG	O5-C5-C6	2.68	111.41	107.20
2	E	1	NAG	C4-C3-C2	2.57	114.79	111.02
2	H	1	NAG	C1-O5-C5	2.56	115.66	112.19
2	C	1	NAG	C1-O5-C5	2.43	115.48	112.19
2	C	1	NAG	C4-C3-C2	2.41	114.56	111.02
2	D	2	NAG	C4-C3-C2	2.41	114.55	111.02
2	K	2	NAG	O5-C5-C6	2.40	110.96	107.20
2	C	2	NAG	C3-C4-C5	2.39	114.51	110.24
2	I	2	NAG	O5-C5-C6	2.36	110.91	107.20
2	C	2	NAG	C2-N2-C7	2.36	126.26	122.90
2	J	2	NAG	C4-C3-C2	2.31	114.41	111.02
2	D	2	NAG	O5-C5-C6	2.31	110.83	107.20
2	E	1	NAG	C1-O5-C5	2.29	115.30	112.19
2	E	2	NAG	C1-C2-N2	2.28	114.38	110.49
2	F	2	NAG	C8-C7-N2	2.23	119.88	116.10
2	I	2	NAG	O5-C1-C2	2.21	114.78	111.29
2	D	1	NAG	C4-C3-C2	2.20	114.24	111.02
2	H	2	NAG	O7-C7-N2	2.17	125.94	121.95
2	E	2	NAG	C4-C3-C2	2.16	114.18	111.02
2	F	1	NAG	C1-O5-C5	2.15	115.11	112.19
2	I	2	NAG	C3-C4-C5	2.08	113.94	110.24
2	K	2	NAG	C2-N2-C7	2.07	125.85	122.90

There are no chirality outliers.

All (22) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	H	2	NAG	C1-C2-N2-C7
2	J	2	NAG	C3-C2-N2-C7
2	E	2	NAG	O5-C5-C6-O6
2	G	2	NAG	O5-C5-C6-O6
2	D	2	NAG	O5-C5-C6-O6
2	I	2	NAG	O5-C5-C6-O6
2	E	2	NAG	C4-C5-C6-O6
2	G	2	NAG	C4-C5-C6-O6
2	H	2	NAG	O5-C5-C6-O6
2	K	1	NAG	C4-C5-C6-O6
2	K	1	NAG	O5-C5-C6-O6
2	D	2	NAG	C4-C5-C6-O6
2	F	2	NAG	C8-C7-N2-C2

*Continued on next page...*

*Continued from previous page...*

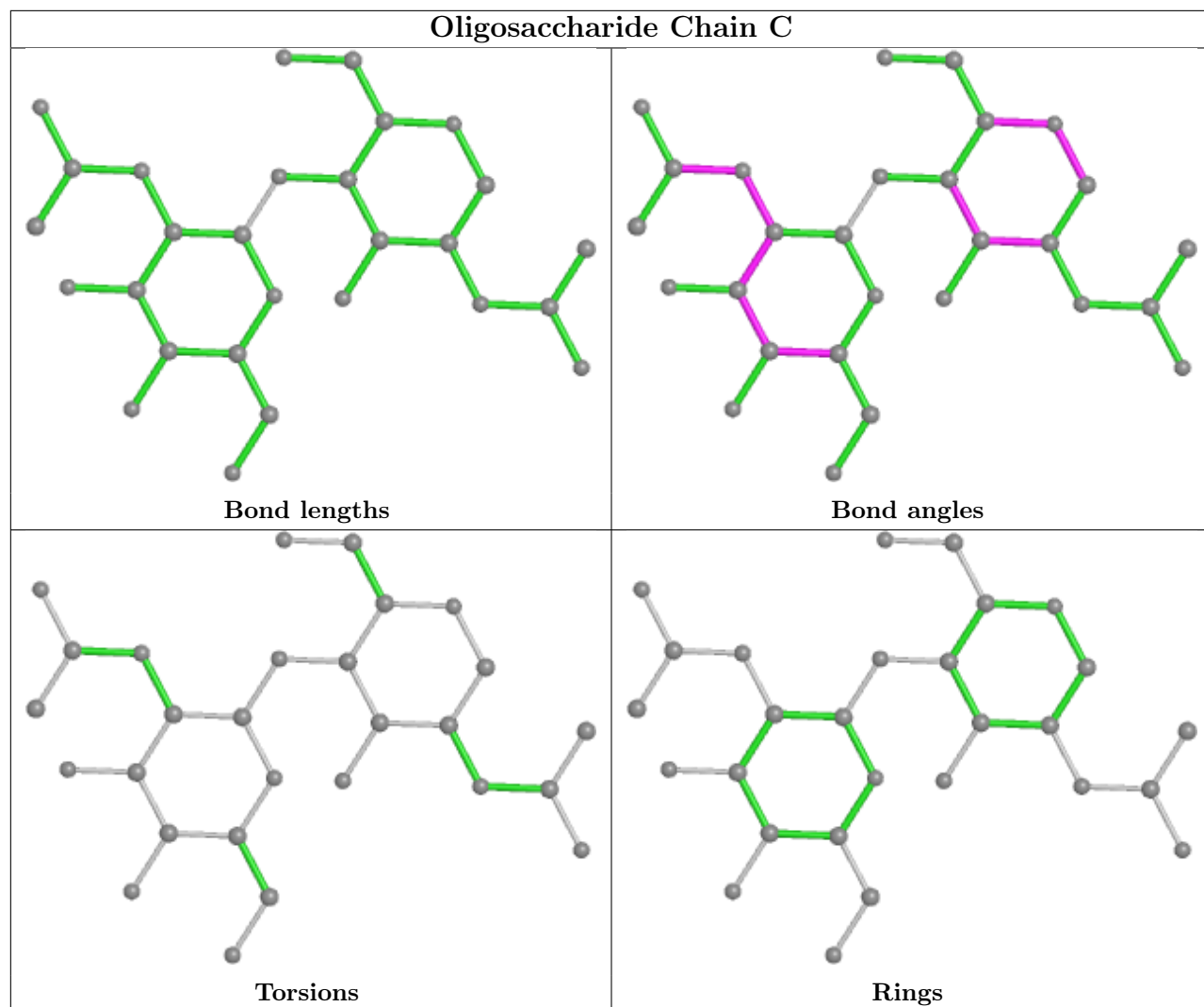
Mol	Chain	Res	Type	Atoms
2	F	2	NAG	O7-C7-N2-C2
2	K	2	NAG	C8-C7-N2-C2
2	K	2	NAG	O7-C7-N2-C2
2	I	2	NAG	C4-C5-C6-O6
2	E	2	NAG	C1-C2-N2-C7
2	K	2	NAG	O5-C5-C6-O6
2	K	2	NAG	C4-C5-C6-O6
2	H	2	NAG	C4-C5-C6-O6
2	F	2	NAG	O5-C5-C6-O6

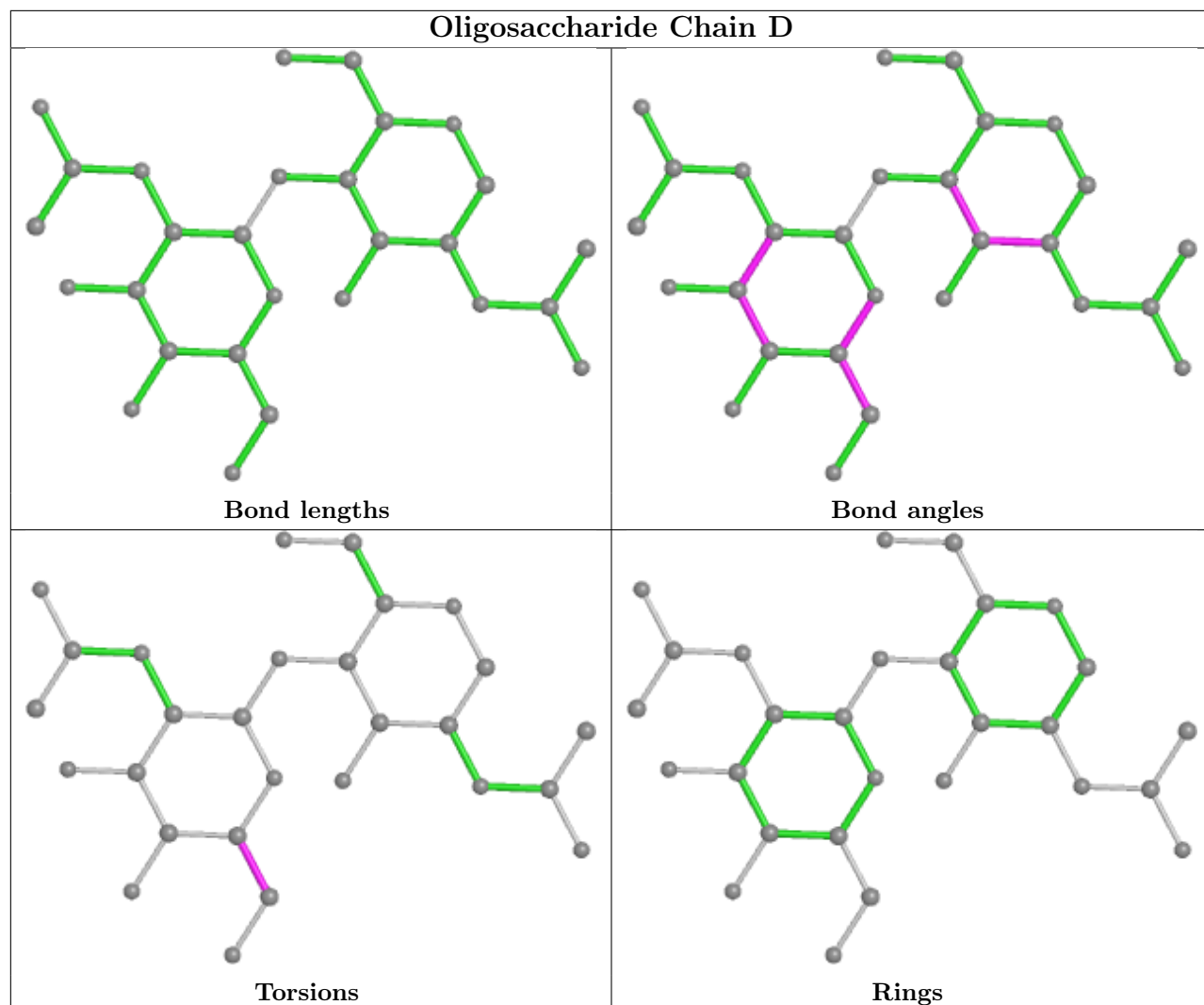
There are no ring outliers.

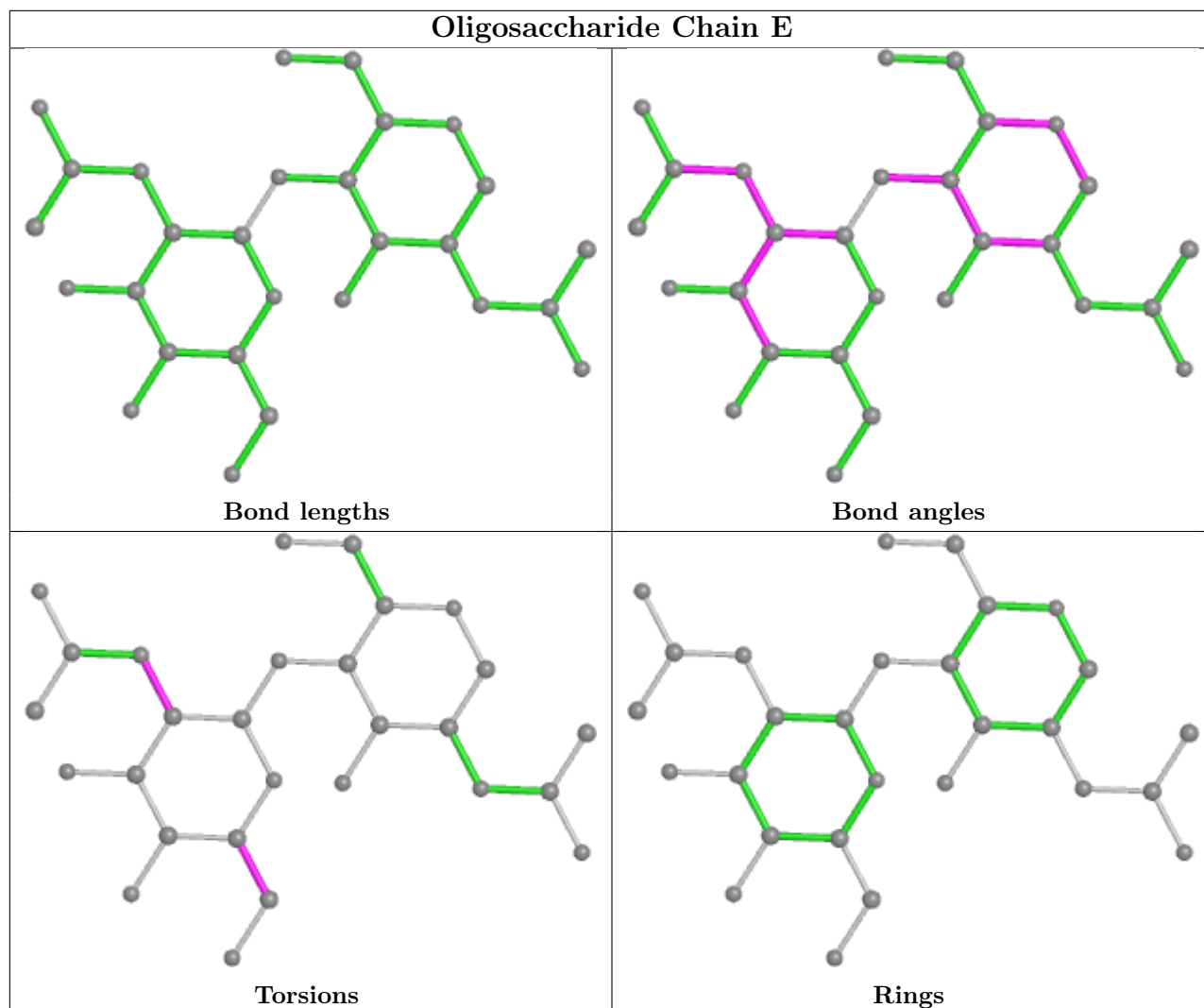
2 monomers are involved in 1 short contact:

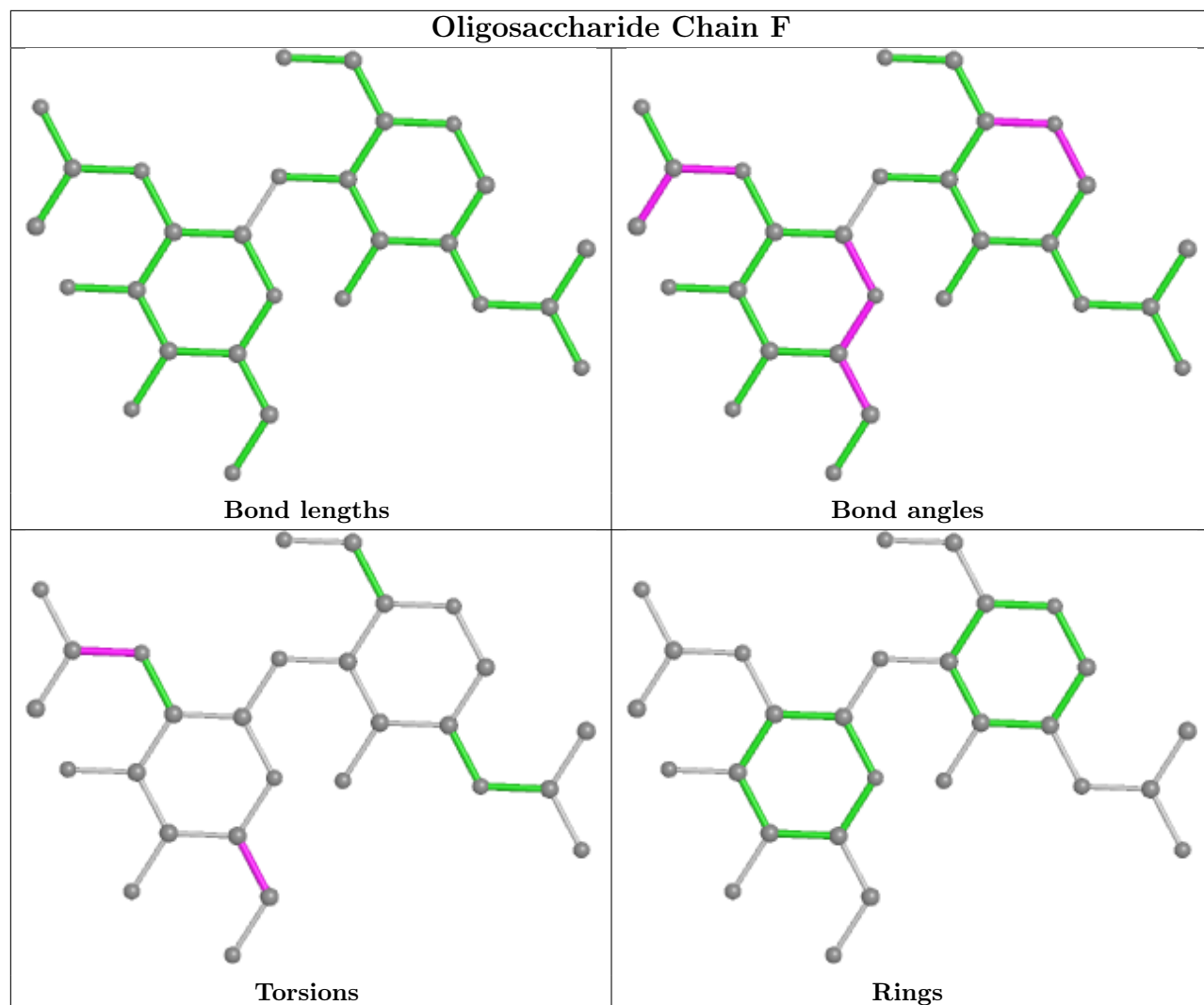
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	J	2	NAG	1	0
2	J	1	NAG	1	0

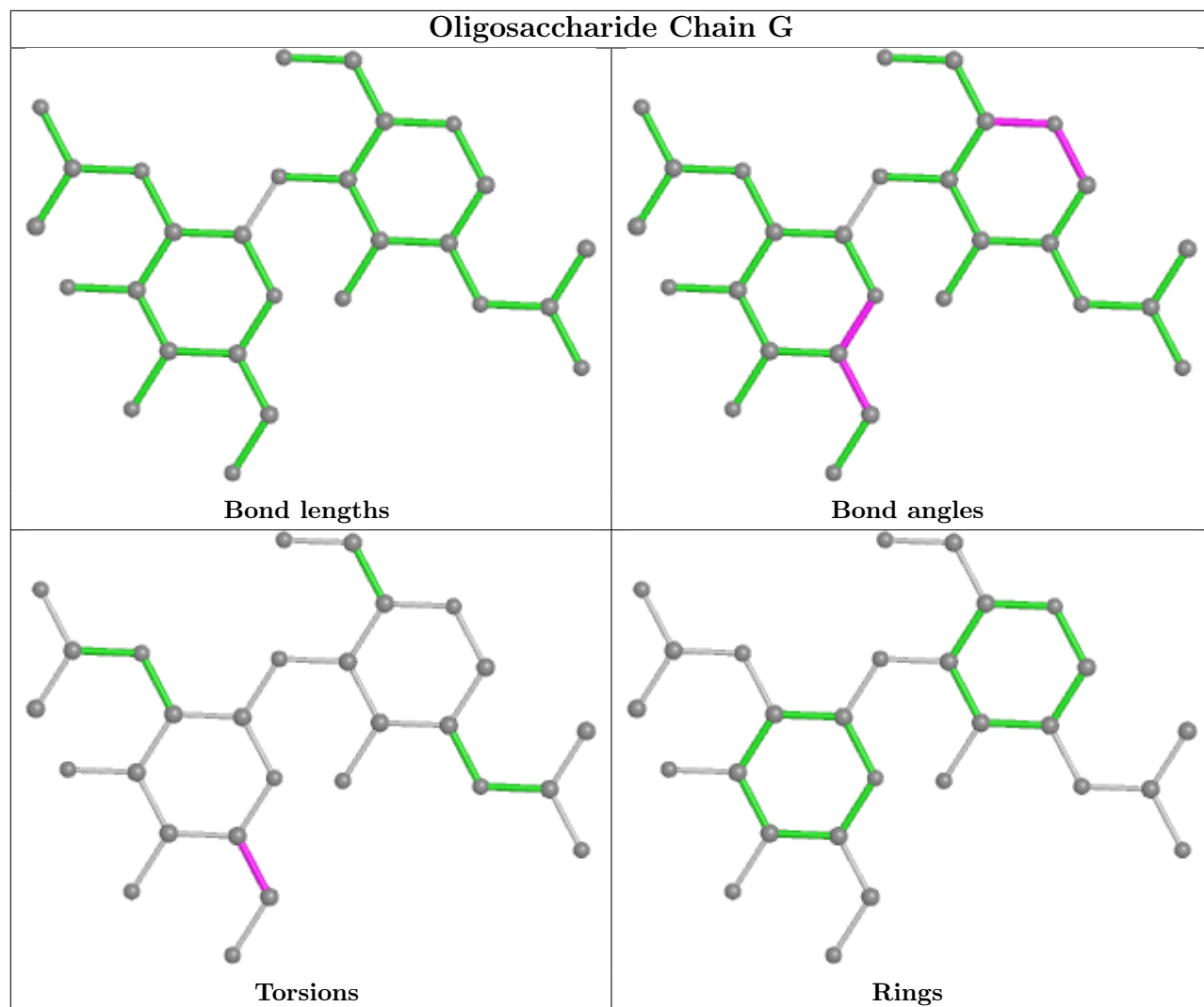
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



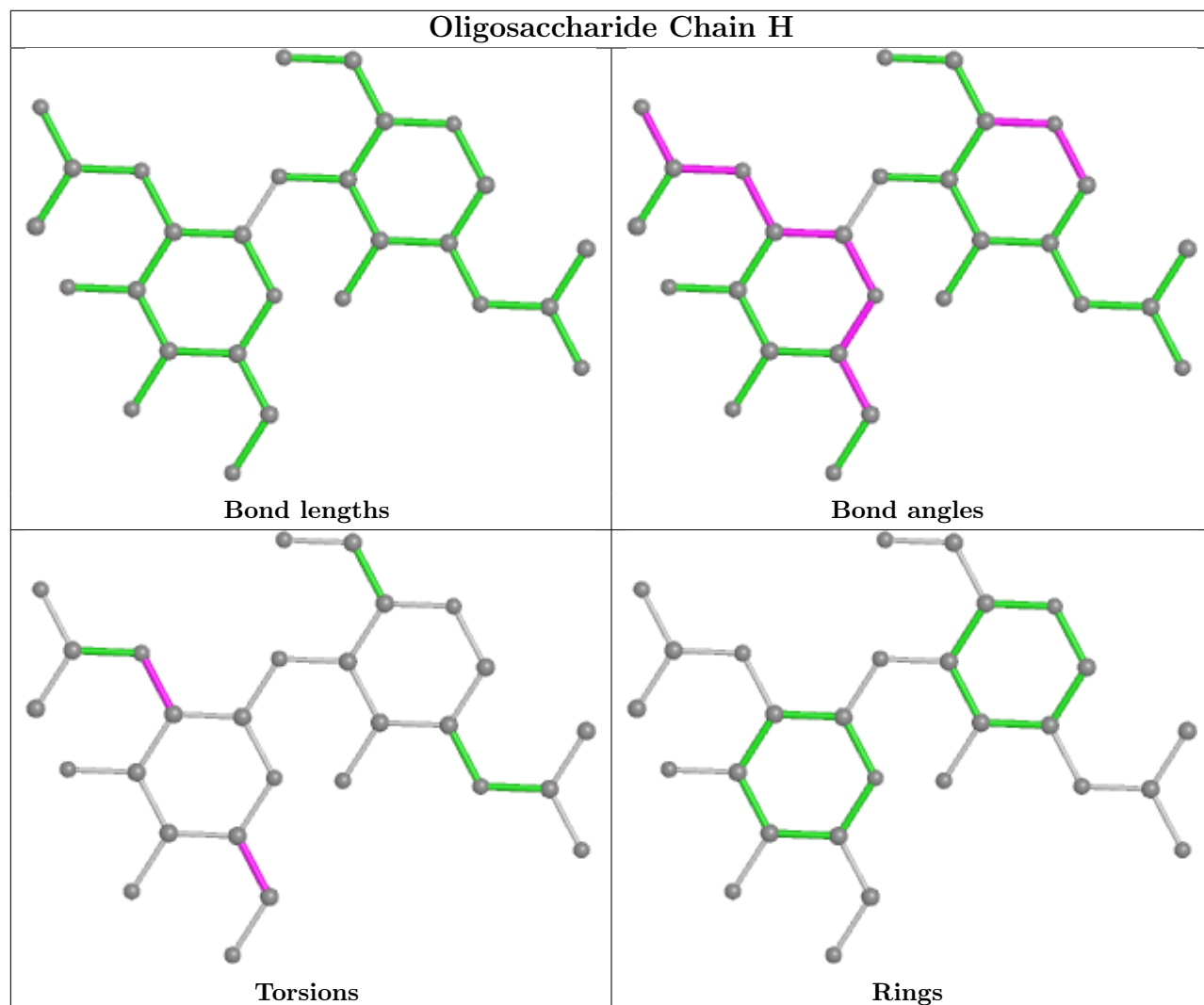


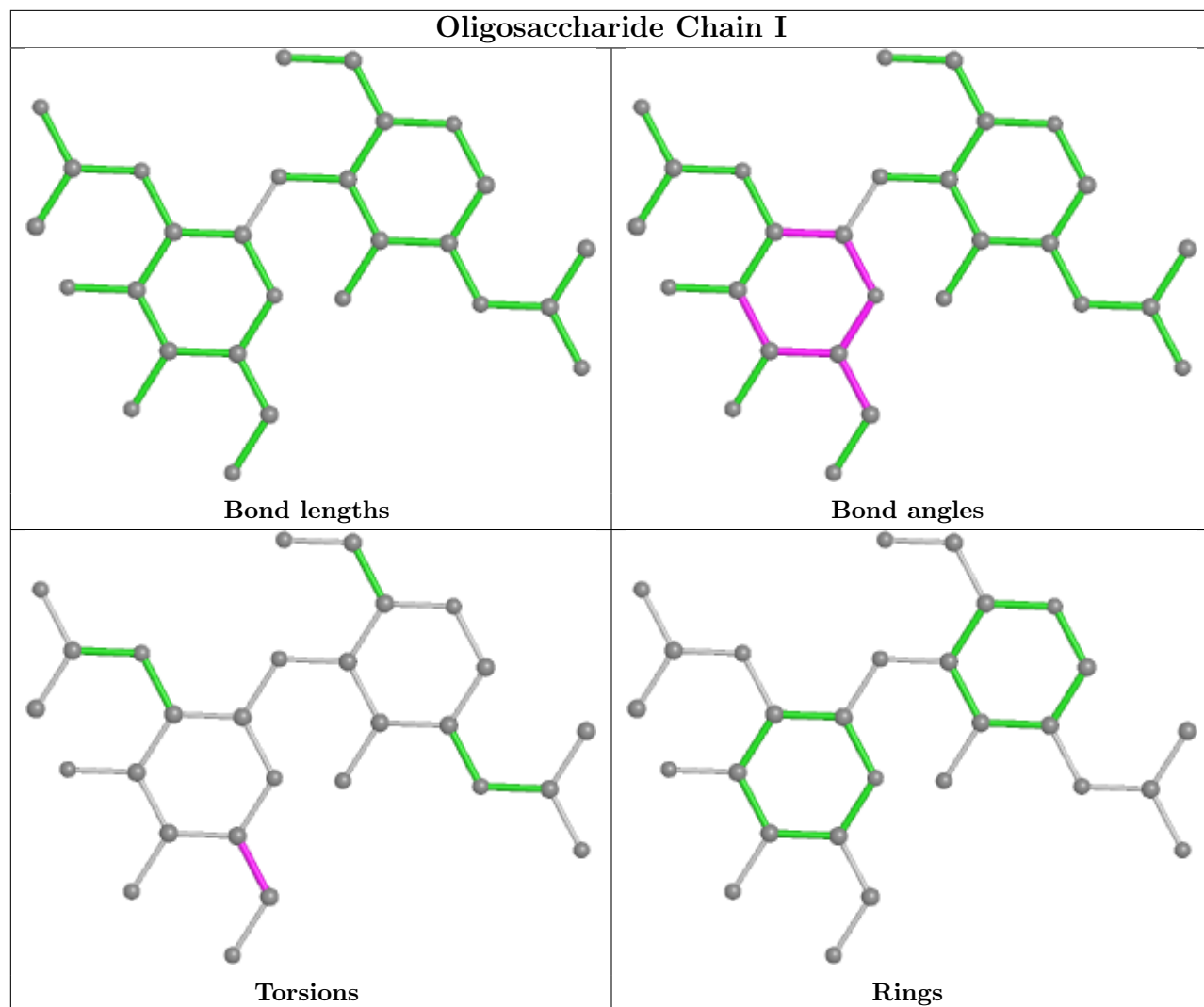


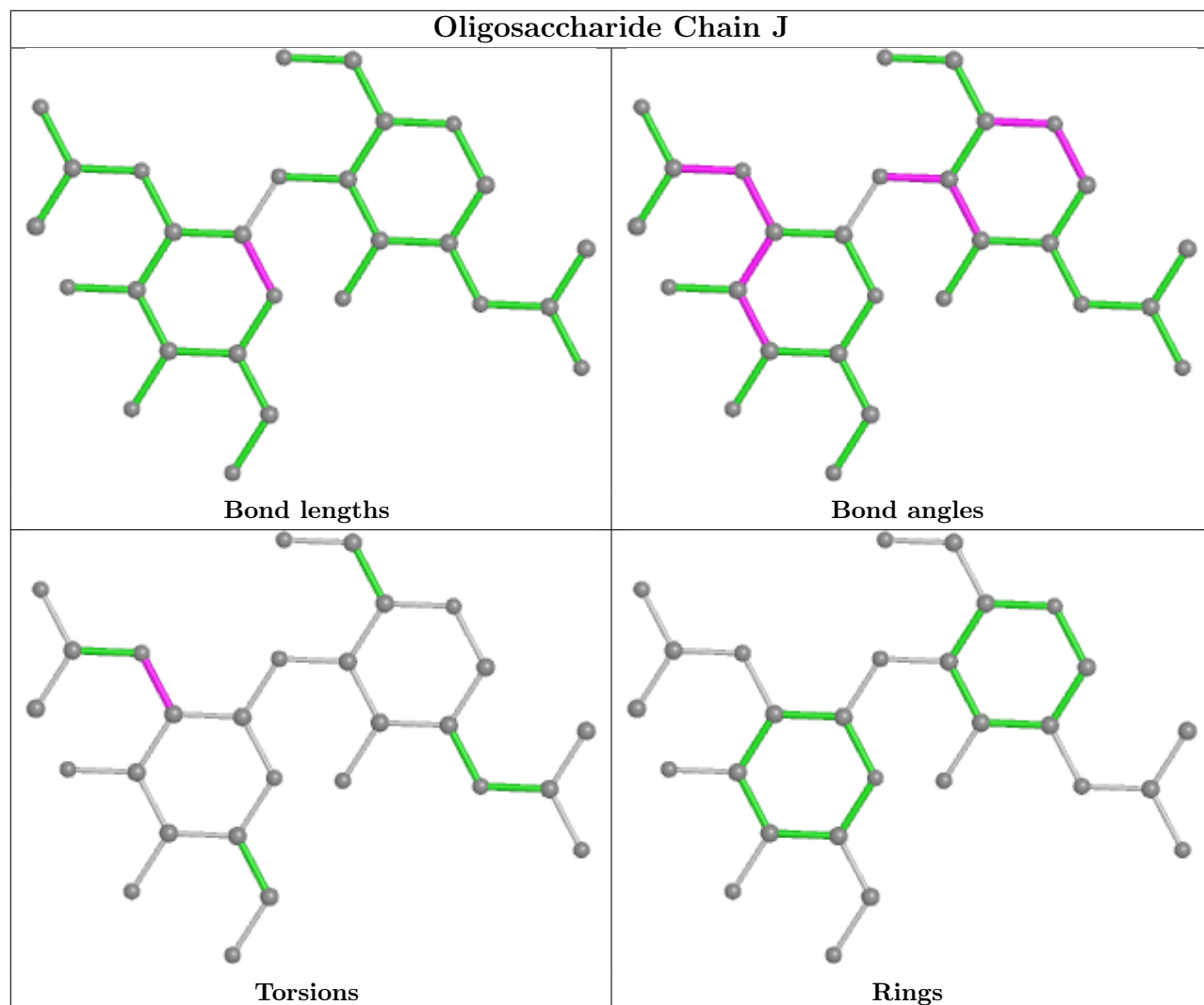


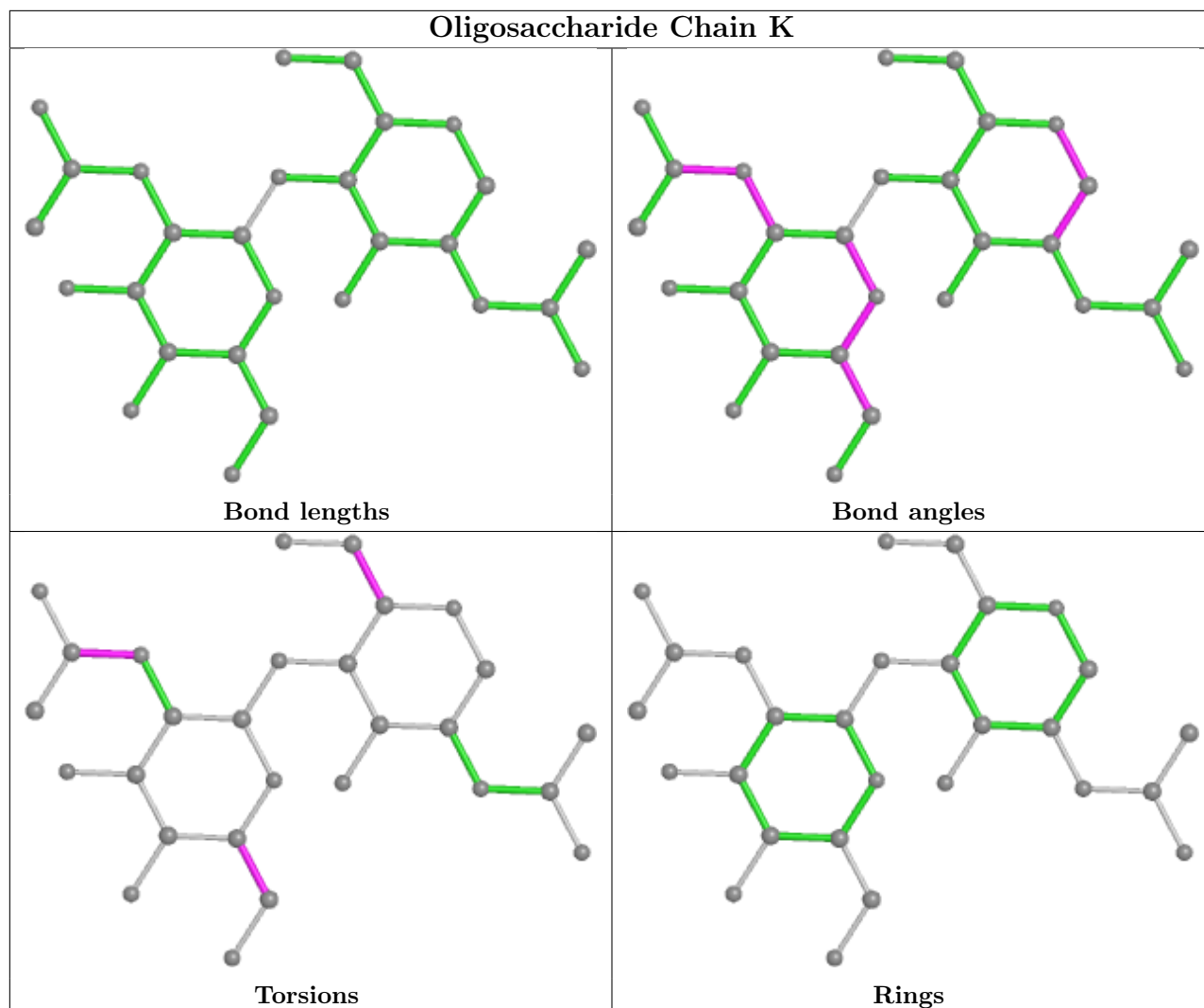












## 5.6 Ligand geometry [i](#)

5 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	NAG	A	808	1	14,14,15	0.60	0	17,19,21	0.97	0
4	C8S	B	812	1	25,28,28	1.58	5 (20%)	41,48,48	1.73	6 (14%)
3	NAG	B	809	1	14,14,15	0.48	0	17,19,21	1.01	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	C8S	A	811	1	25,28,28	1.45	5 (20%)	41,48,48	1.75	7 (17%)
3	NAG	A	801	1	14,14,15	0.50	0	17,19,21	1.26	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	A	808	1	-	0/6/23/26	0/1/1/1
4	C8S	B	812	1	-	2/16/66/66	0/6/5/5
3	NAG	B	809	1	-	0/6/23/26	0/1/1/1
4	C8S	A	811	1	-	2/16/66/66	0/6/5/5
3	NAG	A	801	1	-	0/6/23/26	0/1/1/1

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	812	C8S	C15-N16	-3.65	1.43	1.47
4	B	812	C8S	C23-N24	-2.99	1.36	1.48
4	A	811	C8S	C23-N24	-2.92	1.36	1.48
4	B	812	C8S	C10-C3	2.83	1.57	1.53
4	A	811	C8S	C10-C3	2.66	1.56	1.53
4	B	812	C8S	C17-C15	-2.36	1.49	1.53
4	A	811	C8S	C15-N16	-2.33	1.44	1.47
4	A	811	C8S	C4-C12	2.15	1.58	1.55
4	B	812	C8S	C8-C4	2.11	1.56	1.54
4	A	811	C8S	C17-C15	-2.09	1.50	1.53

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	812	C8S	C20-C19-C17	6.86	132.00	118.37
4	A	811	C8S	C20-C19-C17	6.16	130.60	118.37
4	A	811	C8S	C17-C15-N16	5.39	109.32	101.84
4	B	812	C8S	C17-C15-N16	4.74	108.42	101.84
3	A	801	NAG	C1-O5-C5	4.50	118.29	112.19
4	B	812	C8S	C20-C18-N16	3.37	122.13	115.11
4	A	811	C8S	C10-C3-C5	3.32	113.23	109.25
4	A	811	C8S	C20-C18-N16	3.12	121.62	115.11
3	B	809	NAG	C1-O5-C5	2.89	116.11	112.19

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	812	C8S	C10-C3-C2	2.82	111.56	109.19
4	A	811	C8S	C8-C9-C11	-2.60	106.13	109.25
4	B	812	C8S	C8-C9-C11	-2.34	106.44	109.25
4	B	812	C8S	C10-C3-C5	2.20	111.89	109.25
4	A	811	C8S	C10-C3-C2	2.13	110.98	109.19
4	A	811	C8S	O13-C3-C5	-2.10	103.65	109.28

There are no chirality outliers.

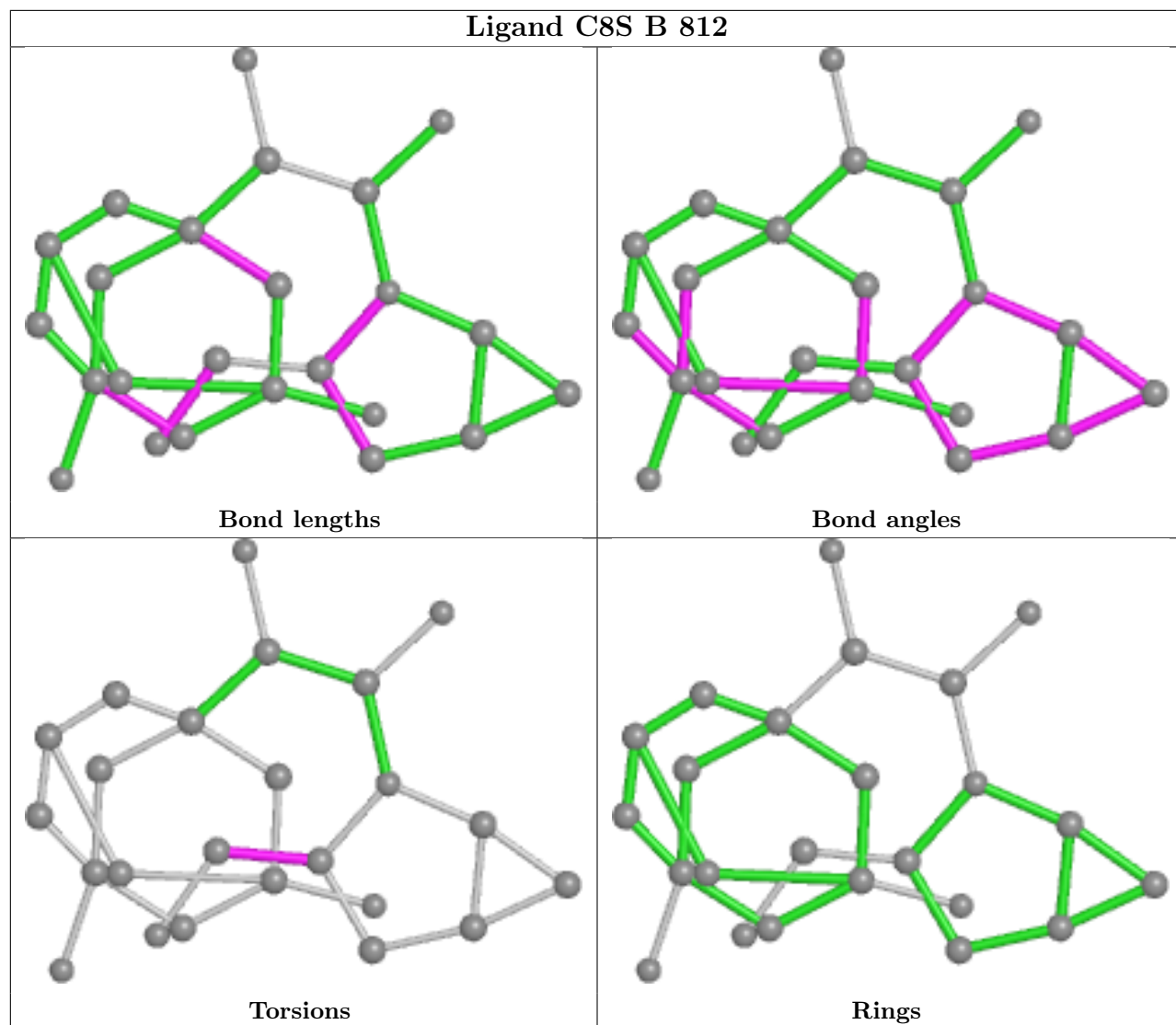
All (4) torsion outliers are listed below:

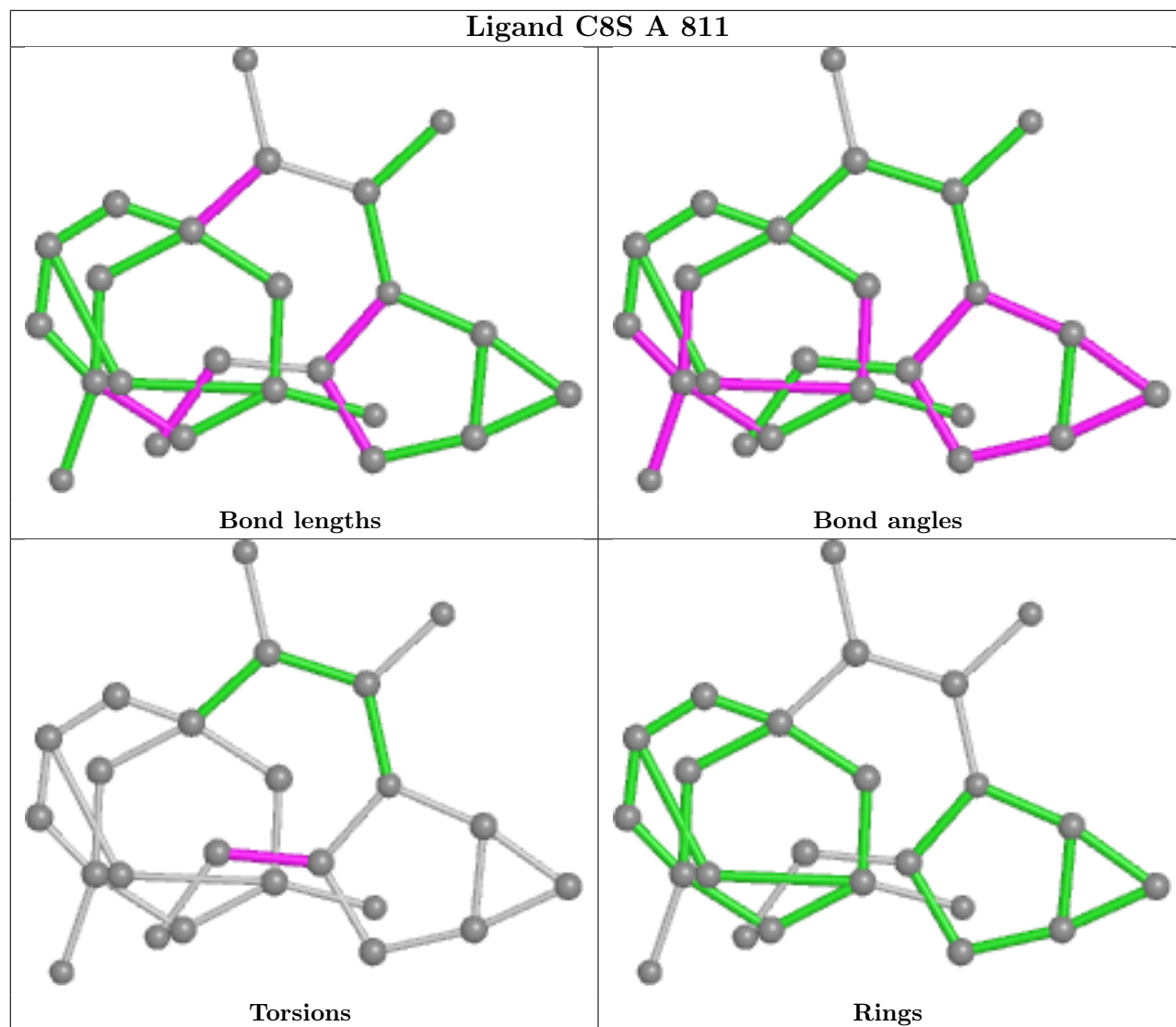
Mol	Chain	Res	Type	Atoms
4	A	811	C8S	N16-C15-C23-N24
4	A	811	C8S	C17-C15-C23-N24
4	B	812	C8S	N16-C15-C23-N24
4	B	812	C8S	C17-C15-C23-N24

There are no ring outliers.

No monomer is involved in short contacts.

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

### 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	728/728 (100%)	0.15	31 (4%) 35 38	15, 24, 46, 66	0
1	B	728/728 (100%)	0.04	23 (3%) 47 50	15, 23, 40, 54	0
All	All	1456/1456 (100%)	0.10	54 (3%) 41 44	15, 24, 43, 66	0

All (54) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	39	THR	9.1
1	A	39	THR	7.8
1	B	766	PRO	6.1
1	A	97	GLU	5.5
1	A	766	PRO	5.1
1	B	279	VAL	5.0
1	B	73	GLU	4.9
1	B	765	LEU	4.6
1	A	140	ARG	4.2
1	B	521	GLU	4.2
1	B	92	ASN	4.0
1	A	40	ARG	3.9
1	B	74	ASN	3.8
1	A	333	SER	3.7
1	A	73	GLU	3.7
1	B	40	ARG	3.6
1	A	99	GLY	3.6
1	B	518	ILE	3.5
1	B	505	GLN	3.3
1	A	96	ASP	3.2
1	A	98	PHE	3.2
1	A	93	SER	3.2
1	A	521	GLU	3.1
1	A	138	ASN	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	141	GLN	2.9
1	A	506	ASN	2.9
1	B	72	GLN	2.9
1	A	74	ASN	2.9
1	A	505	GLN	2.8
1	B	97	GLU	2.8
1	A	91	GLU	2.8
1	A	139	LYS	2.8
1	A	135	TYR	2.8
1	A	334	SER	2.8
1	B	519	LEU	2.7
1	A	378	GLU	2.7
1	B	392	LYS	2.6
1	B	96	ASP	2.6
1	B	93	SER	2.5
1	B	105	TYR	2.5
1	B	520	ASN	2.4
1	B	278	SER	2.3
1	B	280	THR	2.2
1	A	685	ASN	2.2
1	A	71	LYS	2.1
1	A	142	LEU	2.1
1	B	506	ASN	2.1
1	A	95	PHE	2.1
1	A	92	ASN	2.1
1	A	538	LYS	2.1
1	B	538	LYS	2.1
1	A	543	LEU	2.1
1	A	702	LEU	2.0
1	A	100	HIS	2.0

## 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](#)

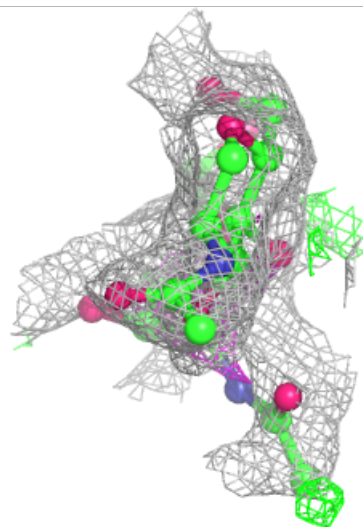
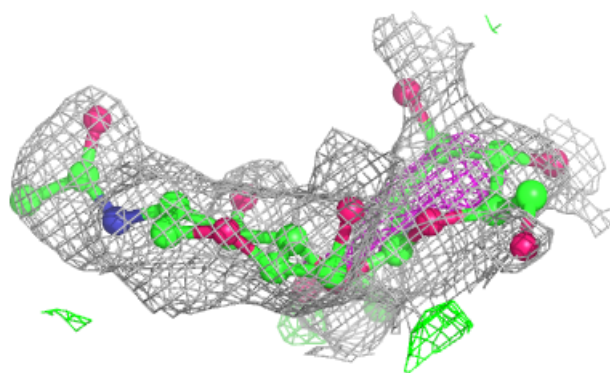
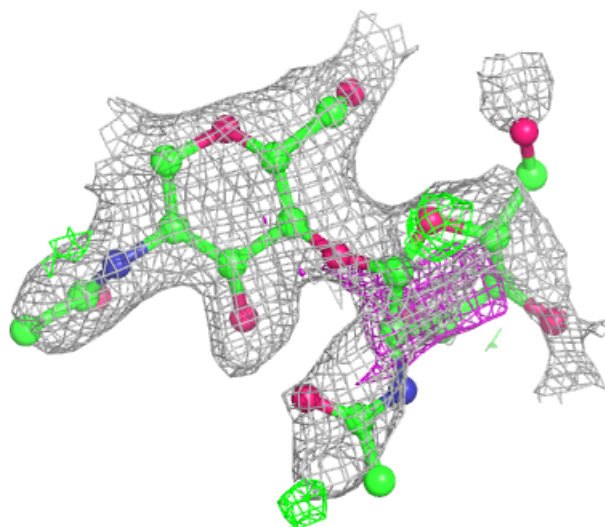
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	H	2	14/15	0.38	0.49	58,61,61,61	0
2	NAG	C	2	14/15	0.54	0.51	70,73,73,73	0
2	NAG	F	2	14/15	0.62	0.55	59,63,65,66	0
2	NAG	J	2	14/15	0.64	0.29	51,54,57,58	0
2	NAG	K	2	14/15	0.64	0.63	65,68,70,70	0
2	NAG	G	2	14/15	0.68	0.44	58,62,63,63	0
2	NAG	D	2	14/15	0.72	0.38	59,62,63,64	0
2	NAG	E	2	14/15	0.73	0.40	54,58,60,60	0
2	NAG	C	1	14/15	0.76	0.30	52,58,61,66	0
2	NAG	H	1	14/15	0.79	0.29	40,45,48,54	0
2	NAG	D	1	14/15	0.83	0.23	43,49,51,55	0
2	NAG	I	2	14/15	0.83	0.31	48,51,54,54	0
2	NAG	K	1	14/15	0.84	0.22	42,49,53,60	0
2	NAG	G	1	14/15	0.86	0.20	41,45,49,54	0
2	NAG	F	1	14/15	0.88	0.24	38,46,48,53	0
2	NAG	E	1	14/15	0.90	0.14	30,33,40,48	0
2	NAG	J	1	14/15	0.92	0.09	30,35,39,46	0
2	NAG	I	1	14/15	0.92	0.15	31,37,40,43	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.

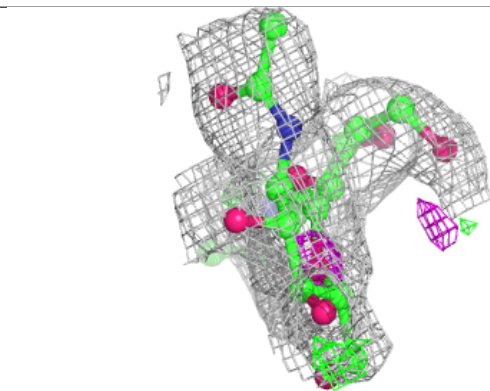
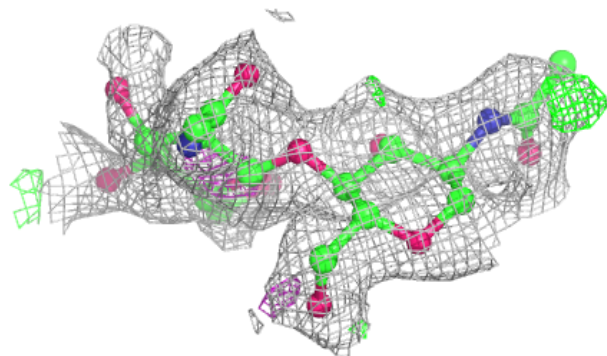
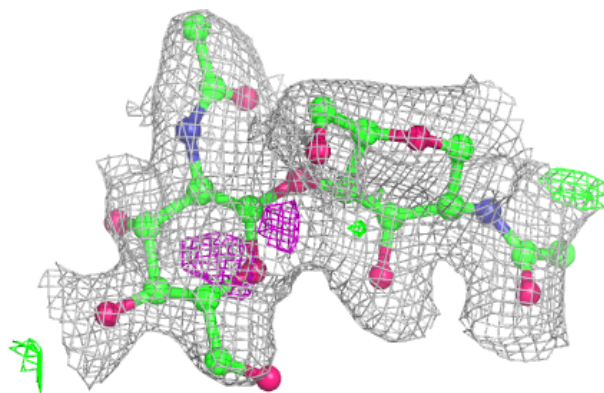
**Electron density around Chain C:**

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

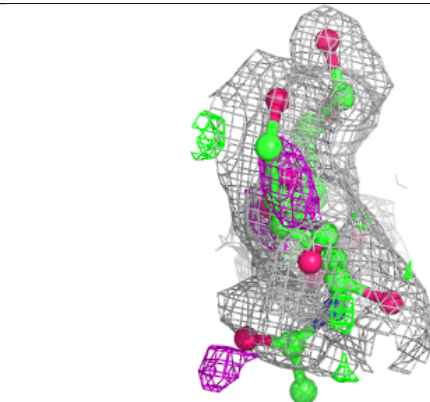
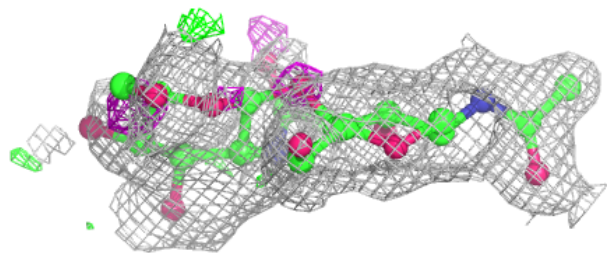
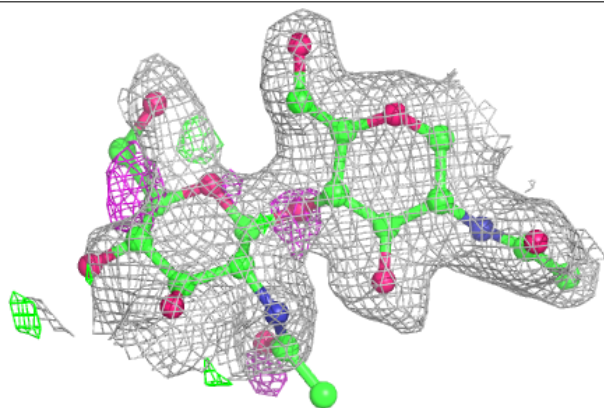


**Electron density around Chain D:**

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

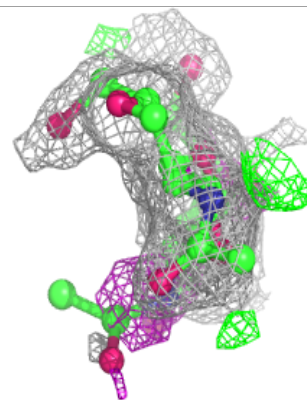
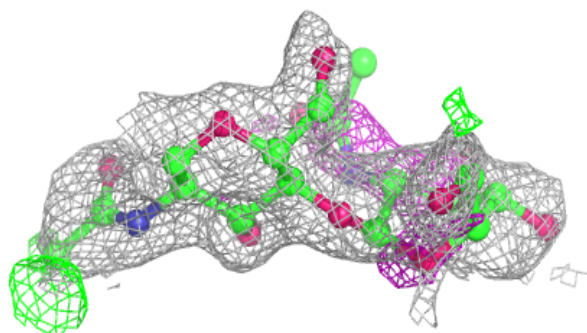
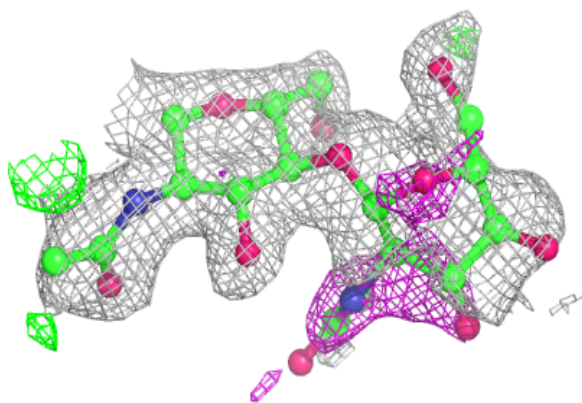
**Electron density around Chain E:**

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



**Electron density around Chain F:**

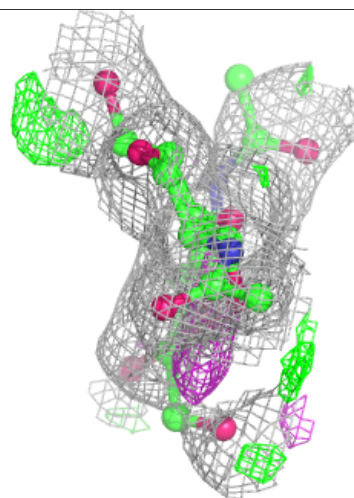
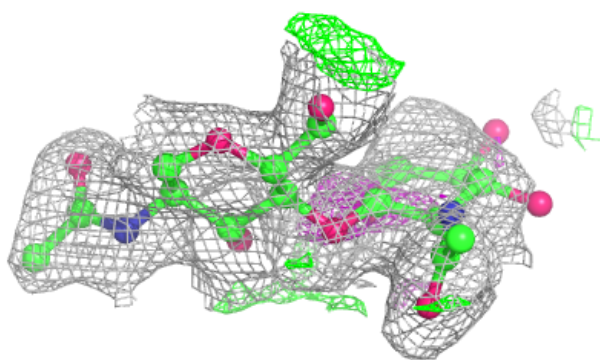
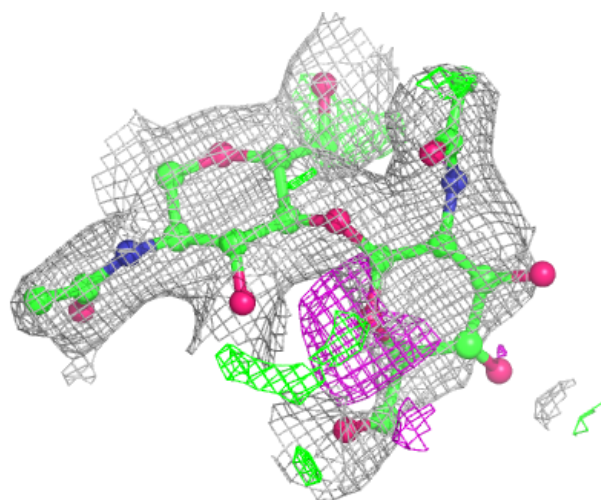
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





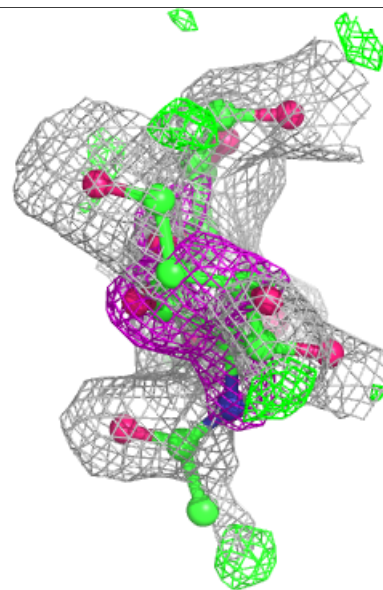
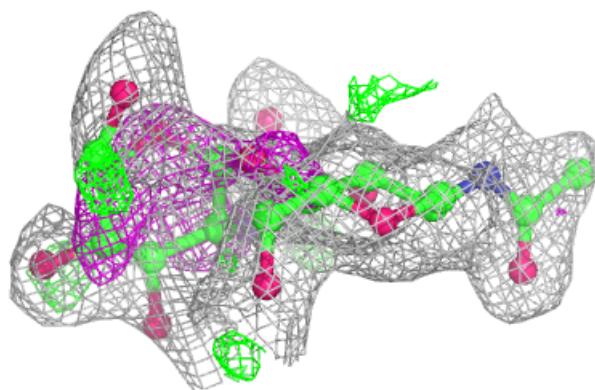
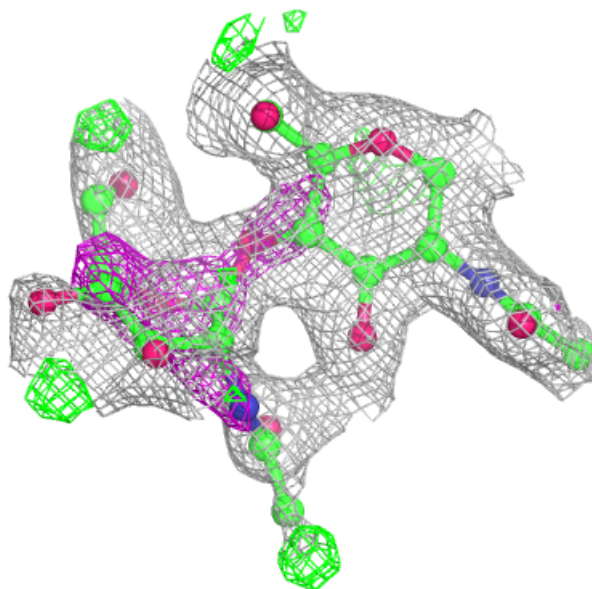
**Electron density around Chain G:**

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



**Electron density around Chain H:**

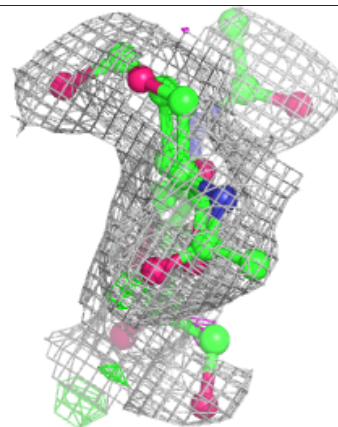
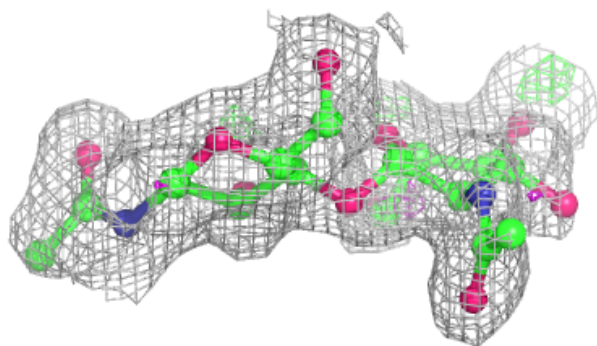
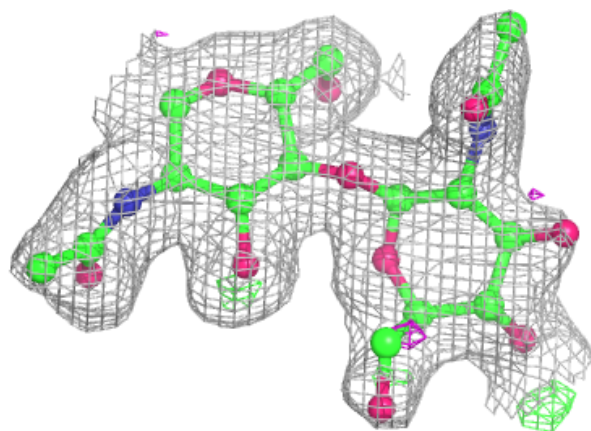
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





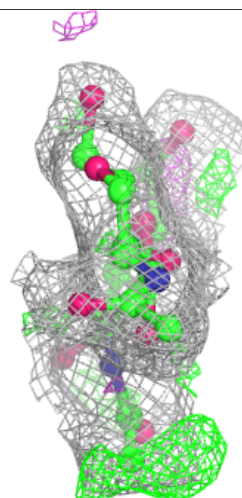
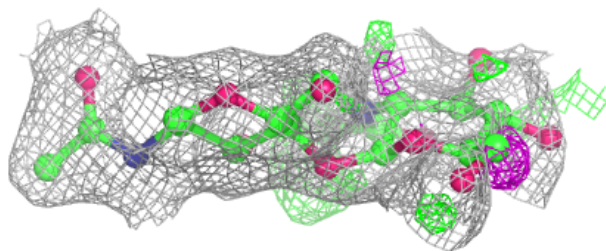
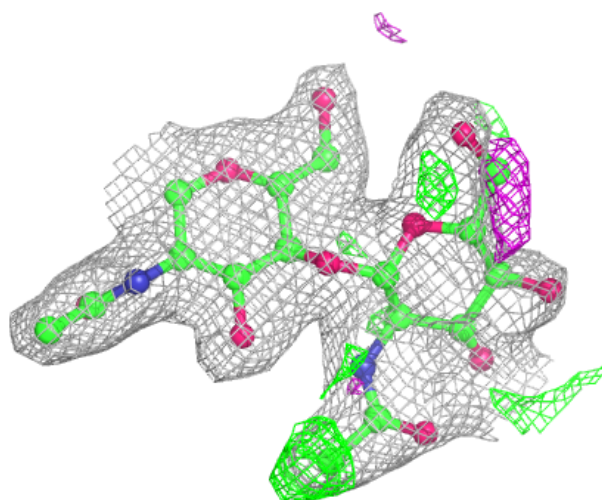
**Electron density around Chain I:**

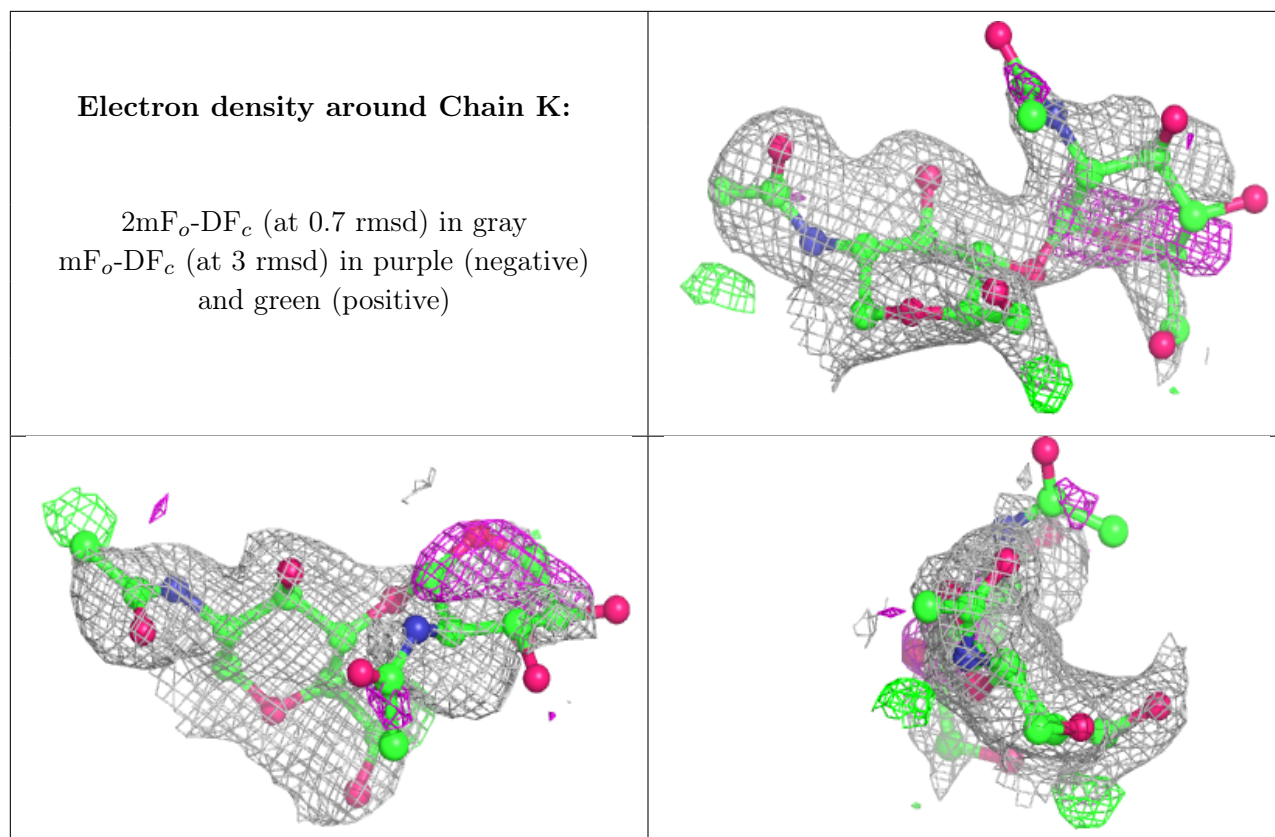
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain J:**

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





## 6.4 Ligands [i](#)

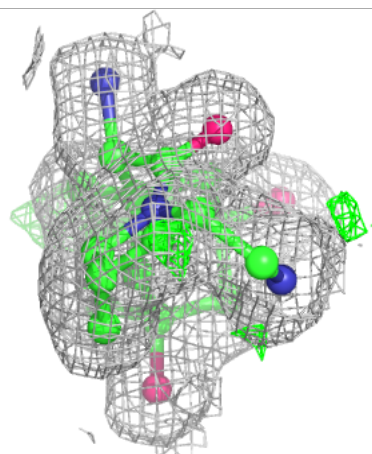
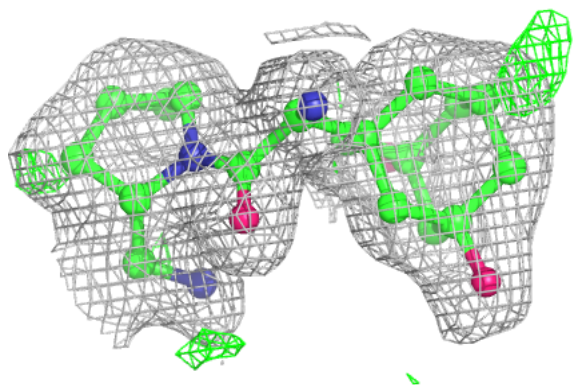
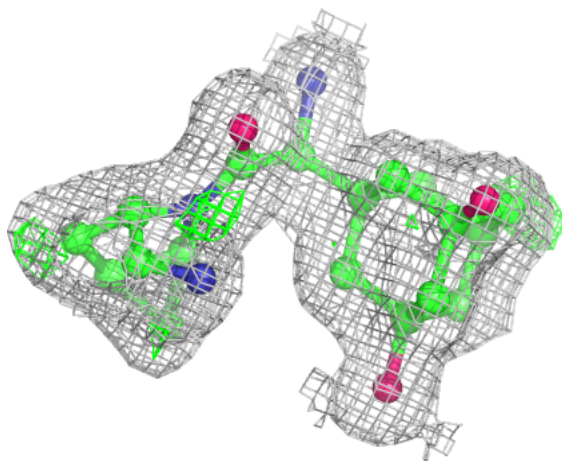
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
3	NAG	B	809	14/15	0.82	0.17	38,42,44,44	0
3	NAG	A	801	14/15	0.86	0.18	51,55,59,63	0
3	NAG	A	808	14/15	0.90	0.11	35,39,41,42	0
4	C8S	A	811	24/24	0.94	0.10	14,16,20,22	0
4	C8S	B	812	24/24	0.95	0.09	15,18,20,22	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.

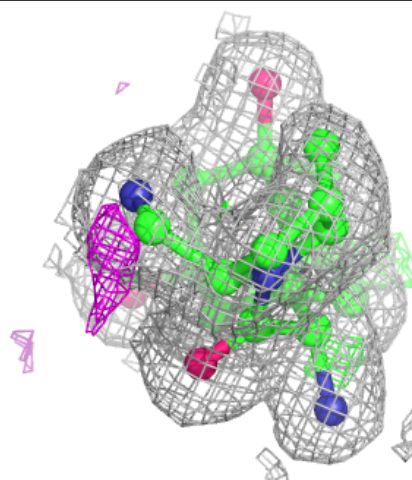
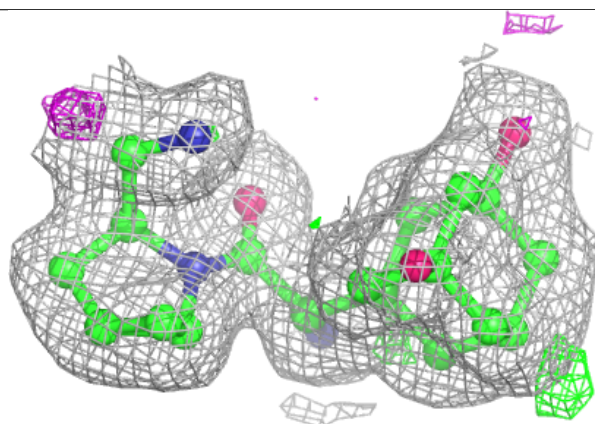
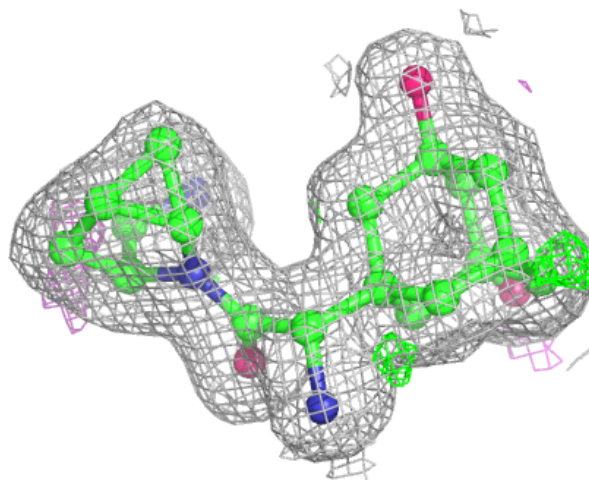
**Electron density around C8S A 811:**

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



**Electron density around C8S B 812:**

$2mF_o-DF_c$  (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.