



# wwPDB X-ray Structure Validation Summary Report

Dec 9, 2023 – 01:49 pm GMT

PDB ID : 1W21  
Title : Structure of Neuraminidase from English duck subtype N6 complexed with 30 mM sialic acid (NANA, Neu5Ac), crystal soaked for 43 hours at 291 K.  
Authors : Rudino-Pinera, E.; Tunnah, P.; Crennell, S.J.; Webster, R.G.; Laver, W.G.; Garman, E.F.  
Deposited on : 2004-06-25  
Resolution : 2.08 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the  symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

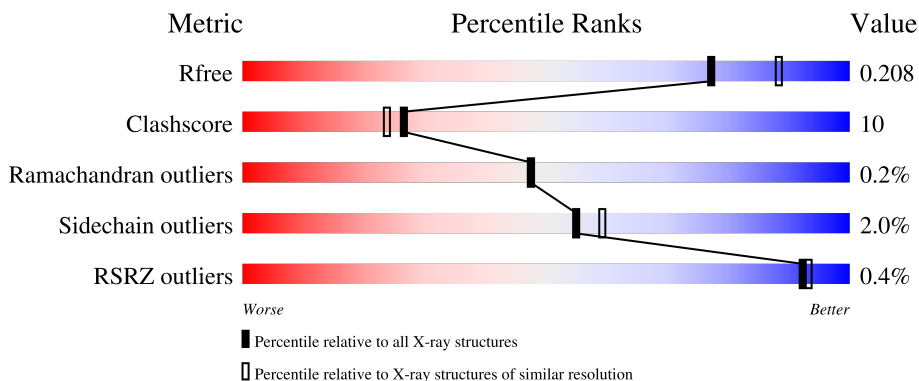
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.08 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.






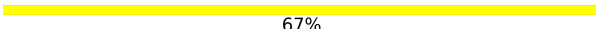

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	6189 (2.10-2.06)
Clashscore	141614	6738 (2.10-2.06)
Ramachandran outliers	138981	6663 (2.10-2.06)
Sidechain outliers	138945	6664 (2.10-2.06)
RSRZ outliers	127900	6057 (2.10-2.06)

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	389	
1	B	389	
1	C	389	
1	D	389	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	E	2	 100%
2	F	2	 100%
2	H	2	 100%
3	G	3	 67%  33%

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	F	1	-	-	X	-
2	NAG	H	2	-	-	X	-
3	BMA	G	2	-	-	X	-
4	SIA	A	1478	-	-	-	X
4	SIA	C	3478	-	-	-	X
5	GOL	A	1479	-	X	-	-
5	GOL	B	2479	-	X	-	-
5	GOL	C	3479	-	X	-	-
5	GOL	D	4479	-	X	-	-
7	BMA	A	1484	-	-	X	-
8	MAN	A	1485	-	-	X	-
8	MAN	C	3486	-	-	-	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 14285 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 NEURAMINIDASE.

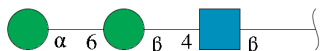
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	389	3008	1875	535	571	27	0	0	0
1	B	389	3008	1875	535	571	27	0	0	0
1	C	389	3008	1875	535	571	27	0	0	0
1	D	389	3009	1875	535	572	27	0	0	0

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



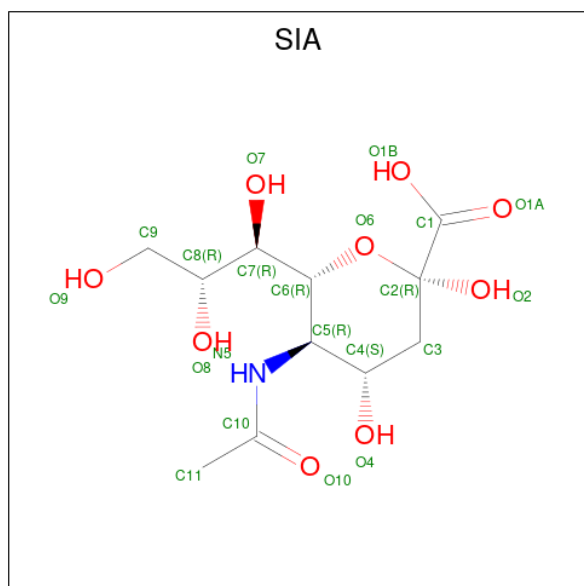
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	E	2	28	16	2	10	0	0	0
2	F	2	28	16	2	10	0	0	0
2	H	2	28	16	2	10	0	0	0

- Molecule 3 is an oligosaccharide called alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	G	3	36	20	1	15	0	0	0

- Molecule 4 is N-acetyl-alpha-neuraminic acid (three-letter code: SIA) (formula:  $C_{11}H_{19}NO_9$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	21	11	1	9	0	0
4	A	1	21	11	1	9	0	0
4	B	1	21	11	1	9	0	0
4	B	1	21	11	1	9	0	0
4	C	1	21	11	1	9	0	0
4	C	1	21	11	1	9	0	0
4	D	1	21	11	1	9	0	0
4	D	1	21	11	1	9	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



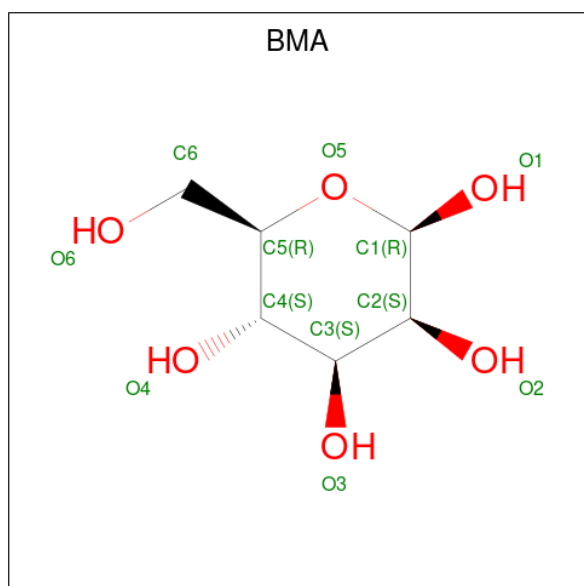
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	C	1	Total	C	O	0	0
			6	3	3		
5	D	1	Total	C	O	0	0
			6	3	3		

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



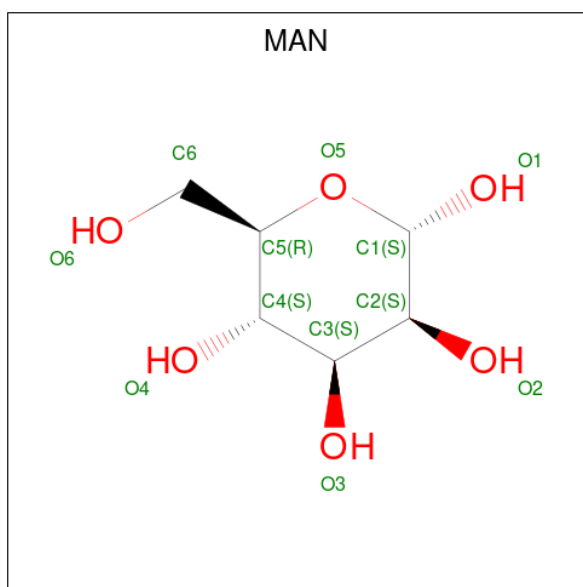
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	B	1	Total	C	N	O	0	0
			14	8	1	5		
6	C	1	Total	C	N	O	0	0
			14	8	1	5		
6	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 7 is beta-D-mannopyranose (three-letter code: BMA) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			11	6	5		
7	B	1	Total	C	O	0	0
			11	6	5		
7	C	1	Total	C	O	0	0
			11	6	5		

- Molecule 8 is alpha-D-mannopyranose (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C O 11 6 5	0	0
8	A	1	Total C O 11 6 5	0	0
8	A	1	Total C O 11 6 5	0	0
8	B	1	Total C O 11 6 5	0	0
8	B	1	Total C O 11 6 5	0	0
8	B	1	Total C O 11 6 5	0	0
8	C	1	Total C O 11 6 5	0	0
8	C	1	Total C O 11 6 5	0	0
8	C	1	Total C O 11 6 5	0	0
8	C	1	Total C O 11 6 5	0	0

- Molecule 9 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1	Total Ca 1 1	0	0
9	B	1	Total Ca 1 1	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	C	1	Total 1	Ca 1	0	0
9	D	1	Total 1	Ca 1	0	0

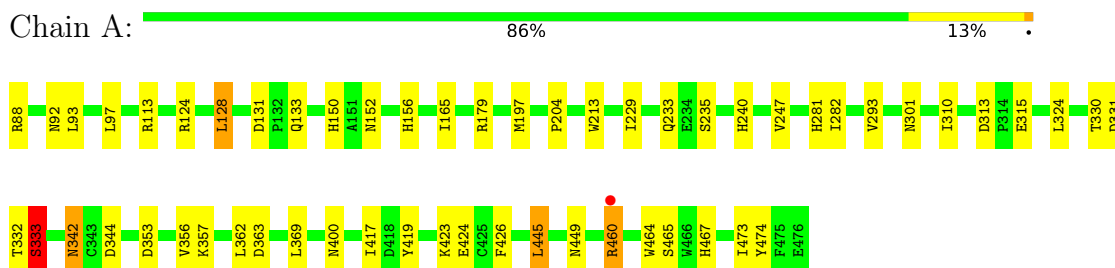
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	429	Total 429	O 429	0	0
10	B	413	Total 413	O 413	0	0
10	C	426	Total 426	O 426	0	0
10	D	455	Total 455	O 455	0	0

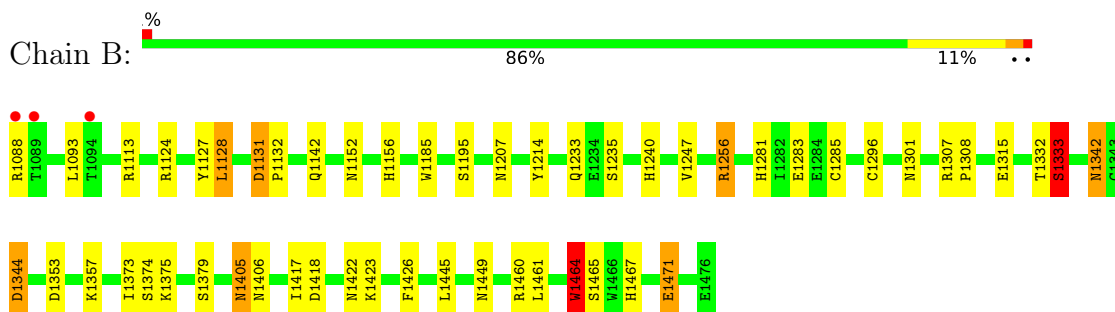
### 3 Residue-property plots i

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

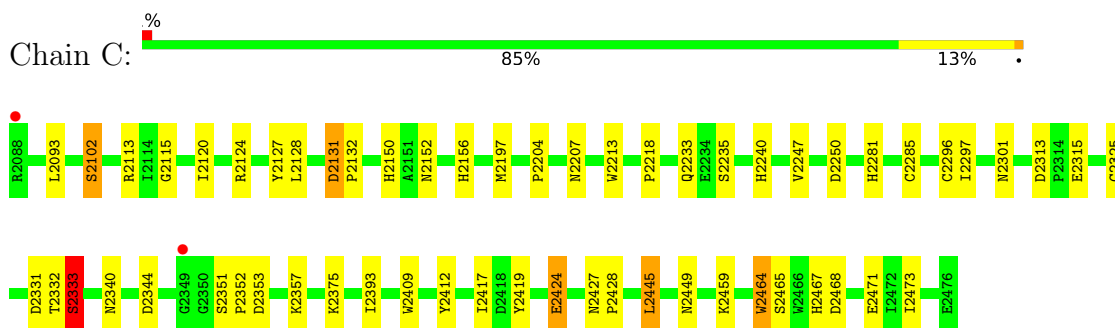
#### • Molecule 1: NEURAMINIDASE



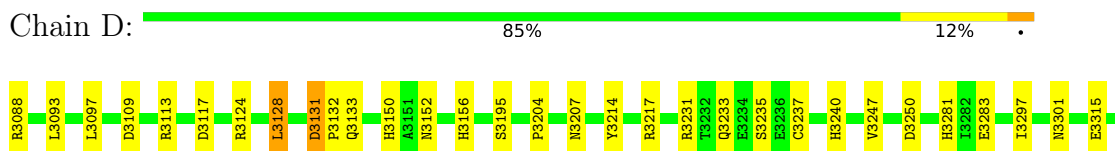
#### • Molecule 1: NEURAMINIDASE



#### • Molecule 1: NEURAMINIDASE



#### • Molecule 1: NEURAMINIDASE





- 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 H:  100%

MAG1  
MAG2

- Molecule 3: alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:  67% 33%

MAG1  
BNA2  
MAN3

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.48Å 74.00Å 106.47Å 90.00° 90.50° 90.00°	Depositor
Resolution (Å)	30.43 – 2.08 29.41 – 2.08	Depositor EDS
% Data completeness (in resolution range)	98.0 (30.43-2.08) 98.0 (29.41-2.08)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.29 (at 2.08Å)	Xtrriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.150 , 0.195 0.161 , 0.208	Depositor DCC
$R_{free}$ test set	4913 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.4	Xtrriage
Anisotropy	1.092	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 60.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.008 for -l,k,h 0.021 for -h,-k,l 0.015 for l,-k,h	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14285	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, GOL, CA, MAN, NAG, SIA

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	1.05	1/3083 (0.0%)	0.89	6/4185 (0.1%)
1	B	1.03	2/3083 (0.1%)	0.89	5/4185 (0.1%)
1	C	1.03	1/3083 (0.0%)	0.90	7/4185 (0.2%)
1	D	1.06	3/3084 (0.1%)	0.93	12/4185 (0.3%)
All	All	1.04	7/12333 (0.1%)	0.90	30/16740 (0.2%)

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	3471	GLU	CD-OE2	6.73	1.33	1.25
1	B	1471	GLU	CD-OE2	6.62	1.32	1.25
1	C	2102	SER	CB-OG	-5.53	1.35	1.42
1	A	474	TYR	CE1-CZ	5.38	1.45	1.38
1	D	3217	ARG	CZ-NH2	5.09	1.39	1.33

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	3250	ASP	CB-CG-OD2	9.23	126.61	118.30
1	B	1344	ASP	CB-CG-OD2	7.47	125.02	118.30
1	B	1333	SER	N-CA-C	-7.36	91.13	111.00
1	C	2131	ASP	CB-CG-OD2	7.34	124.91	118.30
1	C	2333	SER	N-CA-C	-7.28	91.34	111.00

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts

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	3008	0	2888	50	0
1	B	3008	0	2888	54	0
1	C	3008	0	2888	51	0
1	D	3009	0	2888	55	0
2	E	28	0	25	10	0
2	F	28	0	26	10	0
2	H	28	0	26	10	0
3	G	36	0	33	10	0
4	A	42	0	36	4	0
4	B	42	0	36	4	0
4	C	42	0	36	4	0
4	D	42	0	36	4	0
5	A	6	0	4	2	0
5	B	6	0	4	1	0
5	C	6	0	4	1	0
5	D	6	0	4	1	0
6	A	28	0	26	5	0
6	B	14	0	13	3	0
6	C	14	0	13	3	0
6	D	14	0	13	3	0
7	A	11	0	10	6	0
7	B	11	0	10	3	0
7	C	11	0	10	4	0
8	A	33	0	30	6	0
8	B	33	0	30	2	0
8	C	44	0	40	10	0
9	A	1	0	0	0	0
9	B	1	0	0	0	0
9	C	1	0	0	0	0
9	D	1	0	0	0	0
10	A	429	0	0	5	0
10	B	413	0	0	13	0
10	C	426	0	0	8	0
10	D	455	0	0	7	0
All	All	14285	0	12017	247	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 10.

The worst 5 of 247 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:A:92:ASN:HD21	6:A:1480:NAG:C1	1.16	1.52
1:B:1152:ASN:HD21	6:B:2480:NAG:C1	1.35	1.38
1:C:2152:ASN:HD21	6:C:3481:NAG:C1	1.37	1.37
1:A:152:ASN:HD21	6:A:1481:NAG:C1	1.46	1.26
1:D:3152:ASN:HD21	6:D:4480:NAG:C1	1.62	1.12

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	387/389 (100%)	368 (95%)	18 (5%)	1 (0%)	41	39
1	B	387/389 (100%)	372 (96%)	14 (4%)	1 (0%)	41	39
1	C	387/389 (100%)	367 (95%)	19 (5%)	1 (0%)	41	39
1	D	387/389 (100%)	372 (96%)	15 (4%)	0	100	100
All	All	1548/1556 (100%)	1479 (96%)	66 (4%)	3 (0%)	47	47

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	333	SER
1	B	1333	SER
1	C	2333	SER

### 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	331/331 (100%)	326 (98%)	5 (2%)	65	69
1	B	331/331 (100%)	321 (97%)	10 (3%)	41	43
1	C	331/331 (100%)	326 (98%)	5 (2%)	65	69
1	D	331/331 (100%)	324 (98%)	7 (2%)	53	57
All	All	1324/1324 (100%)	1297 (98%)	27 (2%)	55	59

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1464	TRP
1	C	2424	GLU
1	D	3424	GLU
1	C	2333	SER
1	C	2445	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 56 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1449	ASN
1	D	3449	ASN
1	C	2270	GLN
1	D	3427	ASN
1	D	3281	HIS

### 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

9 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	E	1	2	14,14,15	0.27	0	17,19,21	1.02	1 (5%)
2	NAG	E	2	2	14,14,15	0.27	0	17,19,21	1.01	1 (5%)
2	NAG	F	1	2	14,14,15	0.26	0	17,19,21	1.00	1 (5%)
2	NAG	F	2	2	14,14,15	0.26	0	17,19,21	1.01	1 (5%)
3	NAG	G	1	3	14,14,15	0.27	0	17,19,21	1.02	1 (5%)
3	BMA	G	2	3	11,11,12	0.27	0	15,15,17	0.64	0
3	MAN	G	3	3	11,11,12	0.26	0	15,15,17	0.64	0
2	NAG	H	1	2	14,14,15	0.27	0	17,19,21	1.02	1 (5%)
2	NAG	H	2	2	14,14,15	0.27	0	17,19,21	1.01	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
2	NAG	E	1	2	-	2/6/23/26	0/1/1/1
2	NAG	E	2	2	-	1/6/23/26	0/1/1/1
2	NAG	F	1	2	-	0/6/23/26	0/1/1/1
2	NAG	F	2	2	-	4/6/23/26	0/1/1/1
3	NAG	G	1	3	-	2/6/23/26	0/1/1/1
3	BMA	G	2	3	-	2/2/19/22	0/1/1/1
3	MAN	G	3	3	-	0/2/19/22	0/1/1/1
2	NAG	H	1	2	-	0/6/23/26	0/1/1/1
2	NAG	H	2	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	1	NAG	C1-C2-N2	-3.22	104.98	110.49
2	H	1	NAG	C1-C2-N2	-3.22	104.99	110.49
2	H	2	NAG	C1-C2-N2	-3.21	105.01	110.49
2	E	2	NAG	C1-C2-N2	-3.21	105.01	110.49
2	E	1	NAG	C1-C2-N2	-3.20	105.03	110.49

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

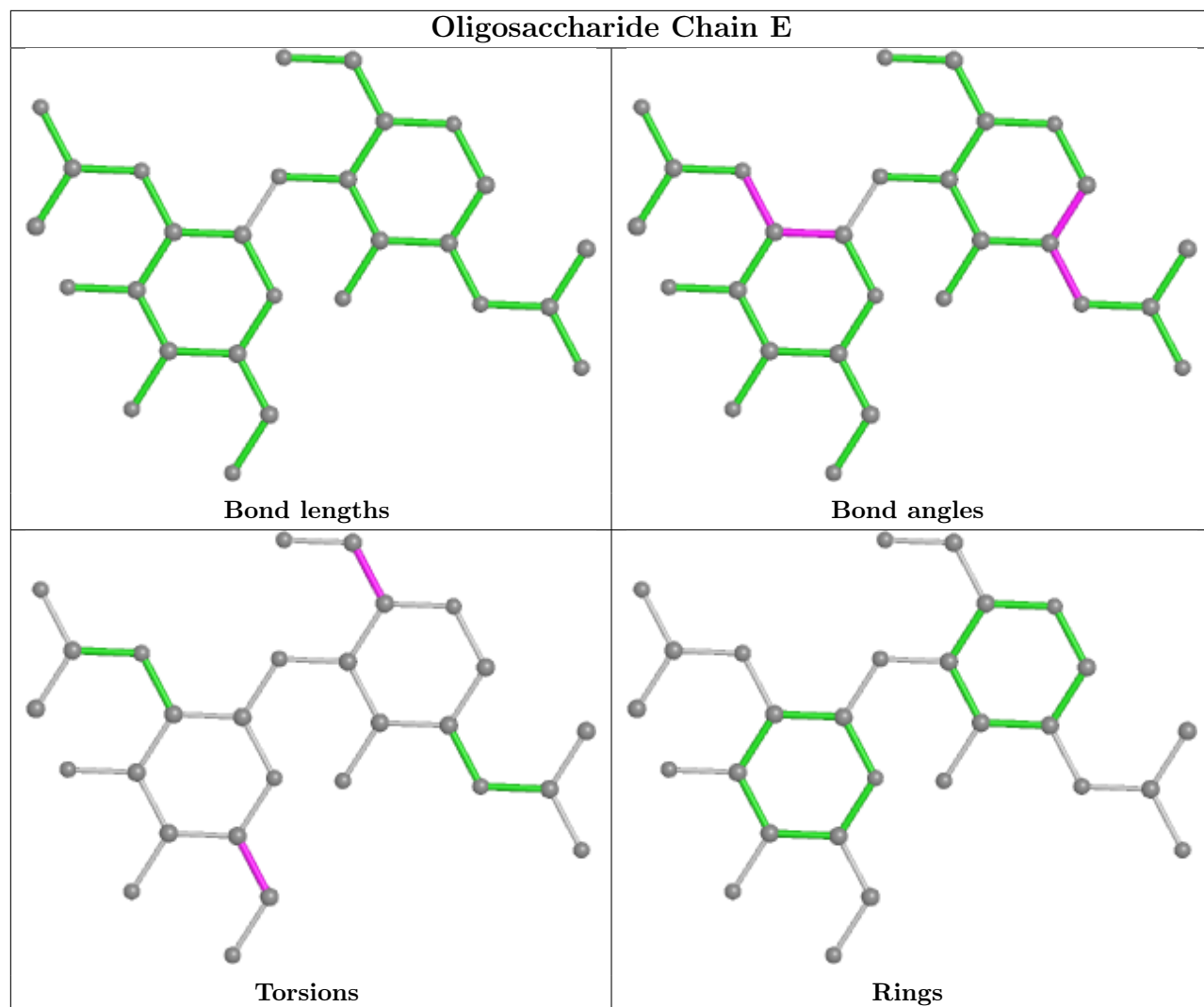
Mol	Chain	Res	Type	Atoms
3	G	1	NAG	C4-C5-C6-O6
2	F	2	NAG	C8-C7-N2-C2
2	F	2	NAG	O7-C7-N2-C2
3	G	1	NAG	O5-C5-C6-O6
2	F	2	NAG	O5-C5-C6-O6

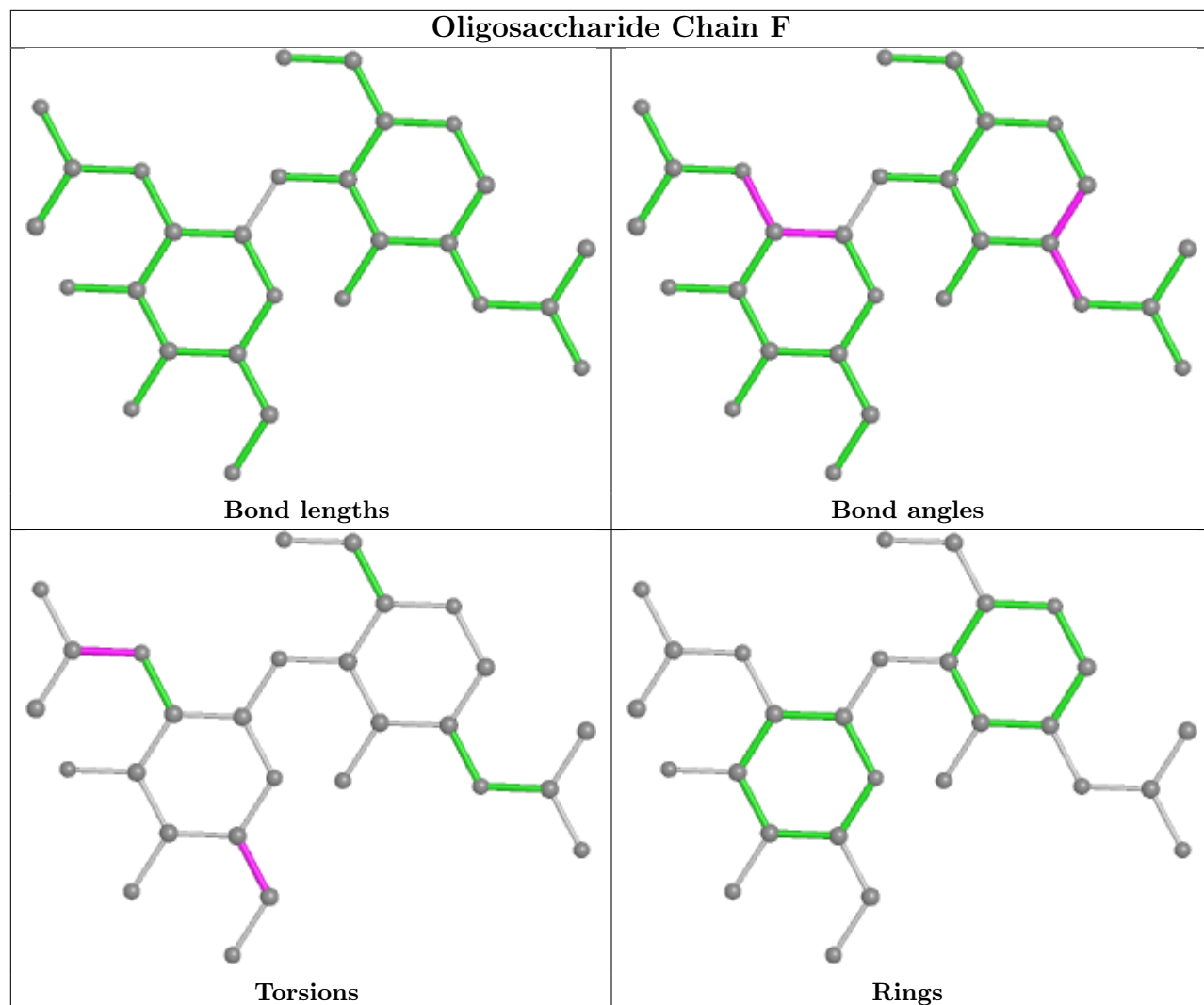
There are no ring outliers.

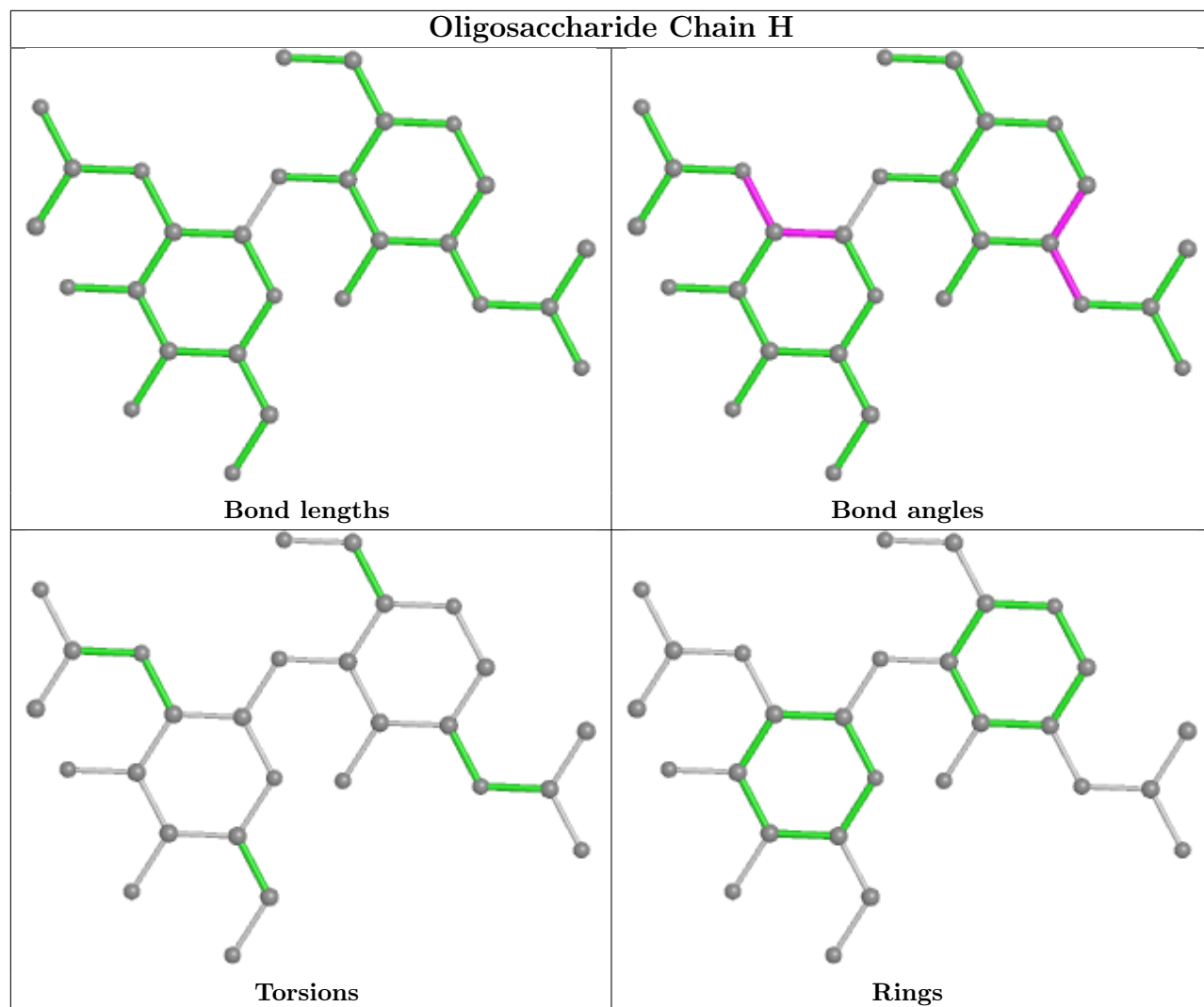
9 monomers are involved in 40 short contacts:

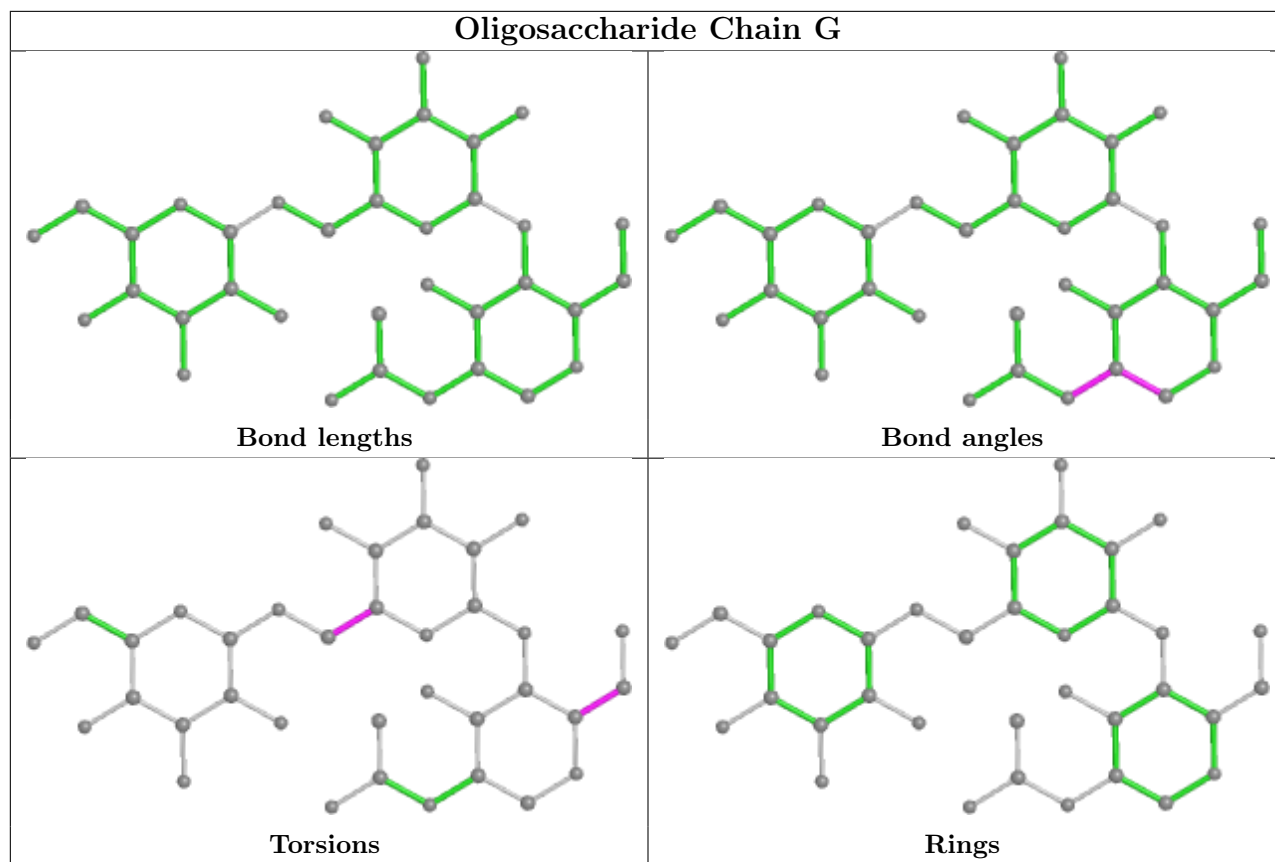
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	1	NAG	2	0
2	E	1	NAG	5	0
2	F	1	NAG	7	0
2	H	1	NAG	5	0
2	F	2	NAG	6	0
2	H	2	NAG	7	0
3	G	3	MAN	4	0
2	E	2	NAG	5	0
3	G	2	BMA	6	0

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 34 ligands modelled in this entry, 4 are monoatomic - leaving 30 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	$\# Z  > 2$
8	MAN	A	1486	-	11,11,12	0.27	0	15,15,17	0.64	0
6	NAG	A	1481	-	14,14,15	0.27	0	17,19,21	1.00	1 (5%)
7	BMA	A	1484	-	11,11,12	0.26	0	15,15,17	0.63	0
6	NAG	C	3481	-	14,14,15	0.28	0	17,19,21	1.01	1 (5%)
7	BMA	C	3490	-	11,11,12	0.26	0	15,15,17	0.64	0
4	SIA	A	1478	-	21,21,21	0.91	1 (4%)	25,31,31	1.29	3 (12%)
8	MAN	A	1485	-	11,11,12	0.26	0	15,15,17	0.64	0
5	GOL	C	3479	-	5,5,5	3.33	3 (60%)	5,5,5	6.32	4 (80%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	A	1480	1	14,14,15	0.26	0	17,19,21	1.02	1 (5%)
8	MAN	B	2485	-	11,11,12	0.26	0	15,15,17	0.63	0
6	NAG	B	2480	-	14,14,15	0.26	0	17,19,21	1.01	1 (5%)
4	SIA	D	4478	-	21,21,21	0.90	1 (4%)	25,31,31	1.29	3 (12%)
4	SIA	B	2477	-	21,21,21	0.89	1 (4%)	25,31,31	1.31	3 (12%)
4	SIA	C	3478	-	21,21,21	0.89	1 (4%)	25,31,31	1.29	3 (12%)
5	GOL	B	2479	-	5,5,5	3.18	2 (40%)	5,5,5	6.21	4 (80%)
4	SIA	B	2478	-	21,21,21	0.92	1 (4%)	25,31,31	1.29	3 (12%)
8	MAN	B	2486	-	11,11,12	0.28	0	15,15,17	0.64	0
5	GOL	A	1479	-	5,5,5	3.60	2 (40%)	5,5,5	6.34	4 (80%)
7	BMA	B	2483	-	11,11,12	0.26	0	15,15,17	0.64	0
8	MAN	C	3485	-	11,11,12	0.26	0	15,15,17	0.63	0
5	GOL	D	4479	-	5,5,5	3.50	2 (40%)	5,5,5	6.50	4 (80%)
8	MAN	B	2484	-	11,11,12	0.27	0	15,15,17	0.63	0
4	SIA	C	3477	-	21,21,21	0.90	1 (4%)	25,31,31	1.30	3 (12%)
4	SIA	D	4477	-	21,21,21	0.89	1 (4%)	25,31,31	1.30	3 (12%)
8	MAN	C	3488	-	11,11,12	0.26	0	15,15,17	0.64	0
8	MAN	A	1487	-	11,11,12	0.26	0	15,15,17	0.63	0
4	SIA	A	1477	-	21,21,21	0.89	1 (4%)	25,31,31	1.30	3 (12%)
6	NAG	D	4480	-	14,14,15	0.25	0	17,19,21	1.01	1 (5%)
8	MAN	C	3486	-	11,11,12	0.26	0	15,15,17	0.64	0
8	MAN	C	3489	-	11,11,12	0.26	0	15,15,17	0.64	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
8	MAN	A	1486	-	-	2/2/19/22	0/1/1/1
6	NAG	A	1481	-	-	1/6/23/26	0/1/1/1
7	BMA	A	1484	-	-	2/2/19/22	0/1/1/1
6	NAG	C	3481	-	-	2/6/23/26	0/1/1/1
7	BMA	C	3490	-	-	0/2/19/22	0/1/1/1
4	SIA	A	1478	-	-	0/20/38/38	0/1/1/1
8	MAN	A	1485	-	-	0/2/19/22	0/1/1/1
5	GOL	C	3479	-	-	3/4/4/4	-
6	NAG	A	1480	1	-	0/6/23/26	0/1/1/1
8	MAN	B	2485	-	-	2/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	B	2480	-	-	2/6/23/26	0/1/1/1
4	SIA	D	4478	-	-	0/20/38/38	0/1/1/1
4	SIA	B	2477	-	-	3/20/38/38	0/1/1/1
4	SIA	C	3478	-	-	4/20/38/38	0/1/1/1
5	GOL	B	2479	-	-	2/4/4/4	-
4	SIA	B	2478	-	-	0/20/38/38	0/1/1/1
8	MAN	B	2486	-	-	0/2/19/22	0/1/1/1
5	GOL	A	1479	-	-	2/4/4/4	-
7	BMA	B	2483	-	-	0/2/19/22	0/1/1/1
8	MAN	C	3485	-	-	1/2/19/22	0/1/1/1
5	GOL	D	4479	-	-	2/4/4/4	-
8	MAN	B	2484	-	-	0/2/19/22	0/1/1/1
4	SIA	C	3477	-	-	2/20/38/38	0/1/1/1
4	SIA	D	4477	-	-	2/20/38/38	0/1/1/1
8	MAN	C	3488	-	-	2/2/19/22	0/1/1/1
8	MAN	A	1487	-	-	1/2/19/22	0/1/1/1
4	SIA	A	1477	-	-	2/20/38/38	0/1/1/1
6	NAG	D	4480	-	-	0/6/23/26	0/1/1/1
8	MAN	C	3486	-	-	0/2/19/22	0/1/1/1
8	MAN	C	3489	-	-	0/2/19/22	0/1/1/1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1479	GOL	O1-C1	-5.53	1.19	1.42
5	A	1479	GOL	C1-C2	-5.45	1.29	1.51
5	D	4479	GOL	O1-C1	-5.36	1.19	1.42
5	D	4479	GOL	C1-C2	-5.21	1.30	1.51
5	C	3479	GOL	C1-C2	-5.03	1.31	1.51

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	4479	GOL	O1-C1-C2	12.10	168.22	110.20
5	A	1479	GOL	O1-C1-C2	11.88	167.15	110.20
5	C	3479	GOL	O1-C1-C2	11.85	167.04	110.20
5	B	2479	GOL	O1-C1-C2	11.37	164.73	110.20
5	B	2479	GOL	O3-C3-C2	6.31	140.44	110.20

There are no chirality outliers.



5 of 37 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	3477	SIA	O1B-C1-C2-O6
4	C	3478	SIA	O1B-C1-C2-O6
5	A	1479	GOL	O1-C1-C2-C3
5	A	1479	GOL	O2-C2-C3-O3
5	B	2479	GOL	O1-C1-C2-C3

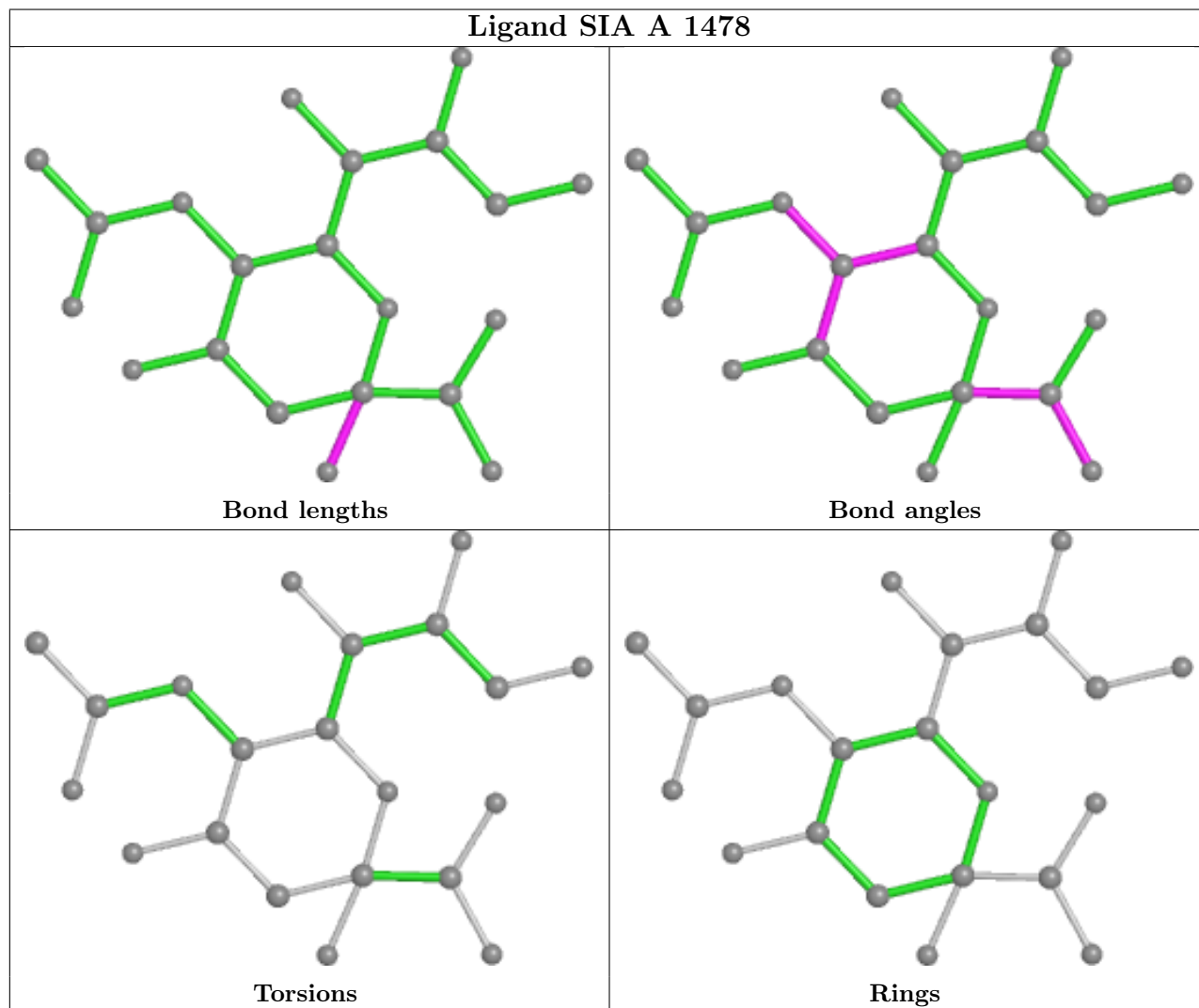
There are no ring outliers.

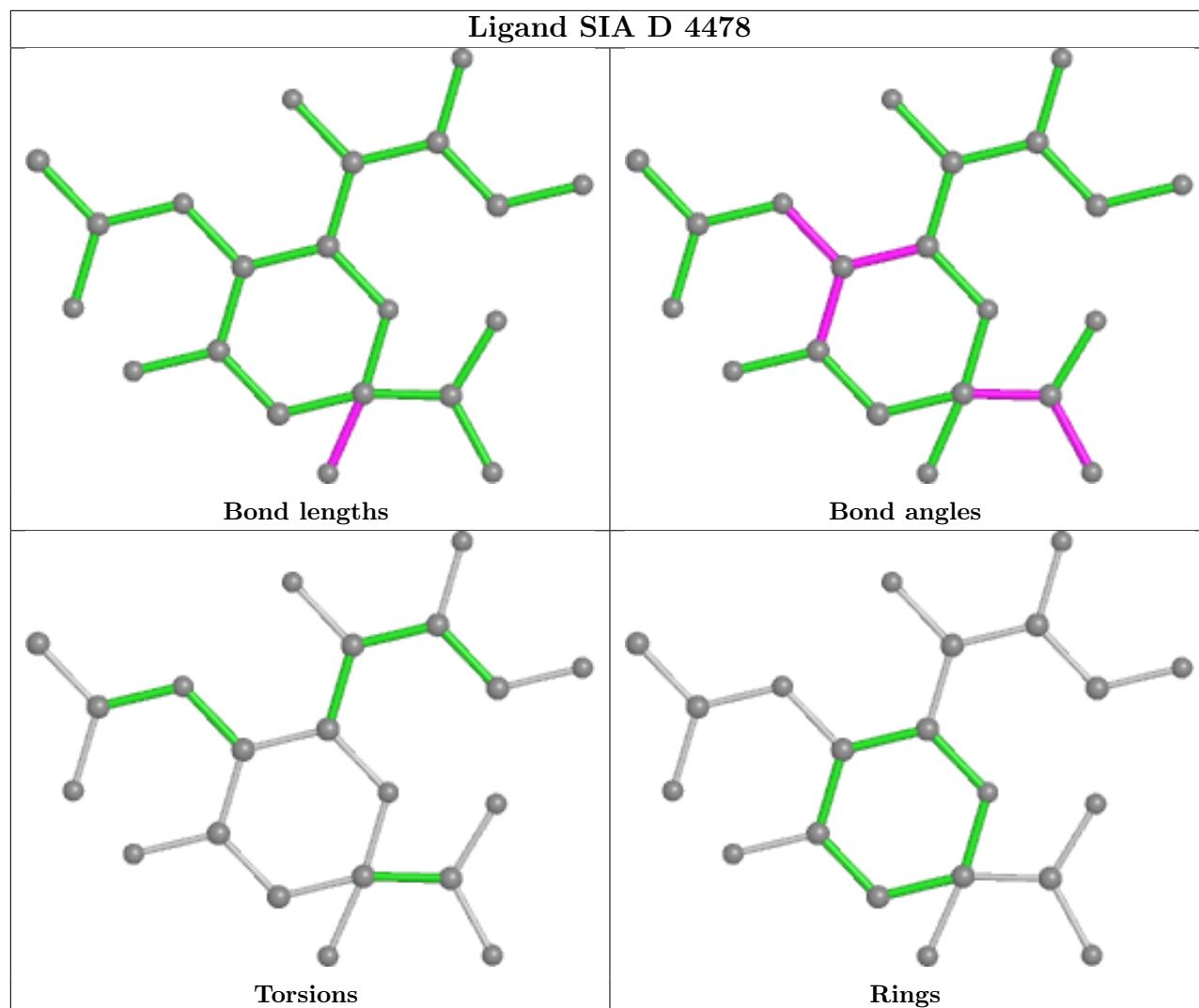
29 monomers are involved in 64 short contacts:

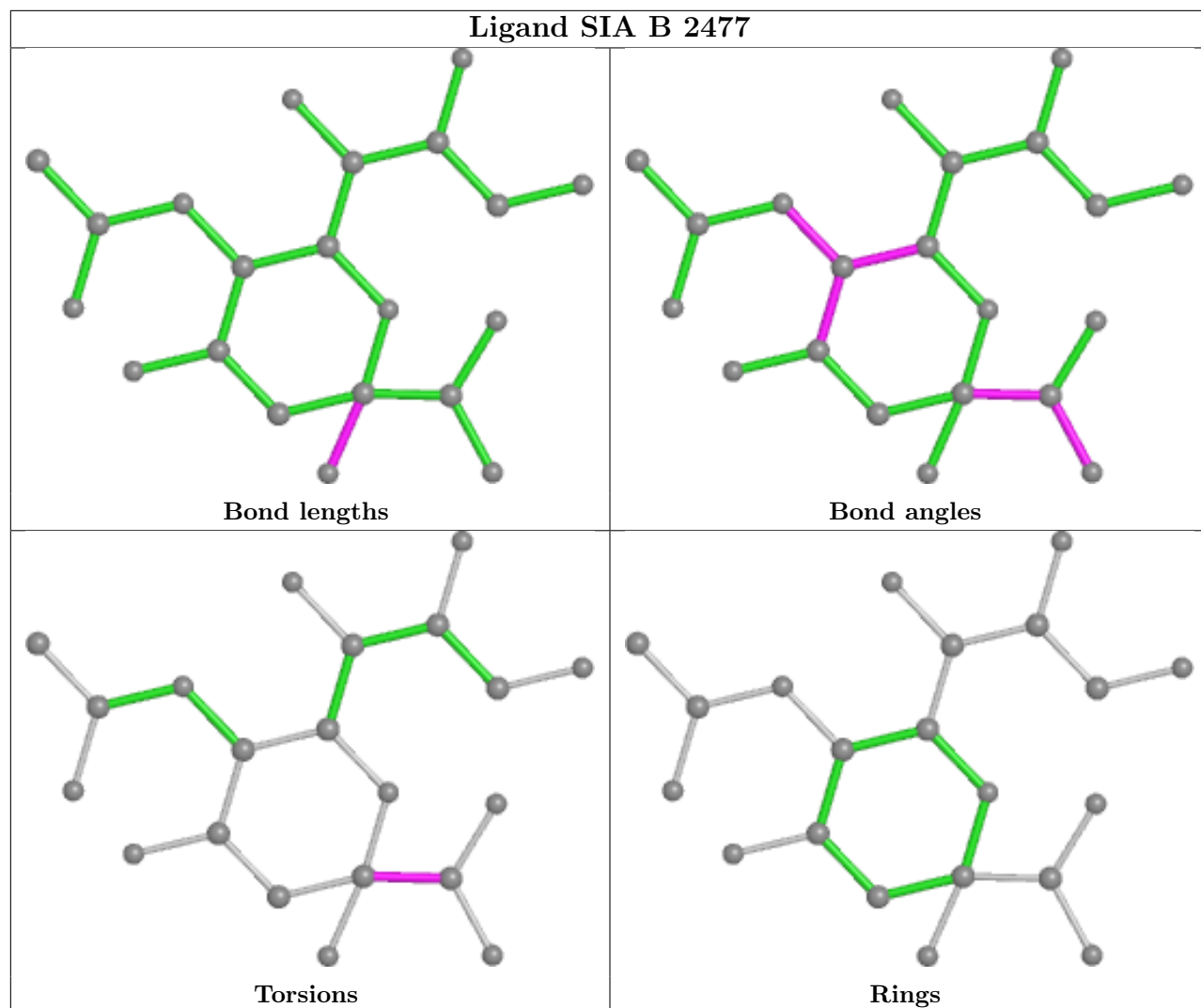
Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	1486	MAN	1	0
6	A	1481	NAG	3	0
7	A	1484	BMA	6	0
6	C	3481	NAG	3	0
7	C	3490	BMA	4	0
4	A	1478	SIA	1	0
8	A	1485	MAN	6	0
5	C	3479	GOL	1	0
6	A	1480	NAG	2	0
8	B	2485	MAN	1	0
6	B	2480	NAG	3	0
4	D	4478	SIA	1	0
4	B	2477	SIA	1	0
4	C	3478	SIA	2	0
5	B	2479	GOL	1	0
4	B	2478	SIA	3	0
5	A	1479	GOL	2	0
7	B	2483	BMA	3	0
8	C	3485	MAN	5	0
5	D	4479	GOL	1	0
8	B	2484	MAN	2	0
4	C	3477	SIA	2	0
4	D	4477	SIA	3	0
8	C	3488	MAN	1	0
8	A	1487	MAN	3	0
4	A	1477	SIA	3	0
6	D	4480	NAG	3	0
8	C	3486	MAN	2	0
8	C	3489	MAN	3	0

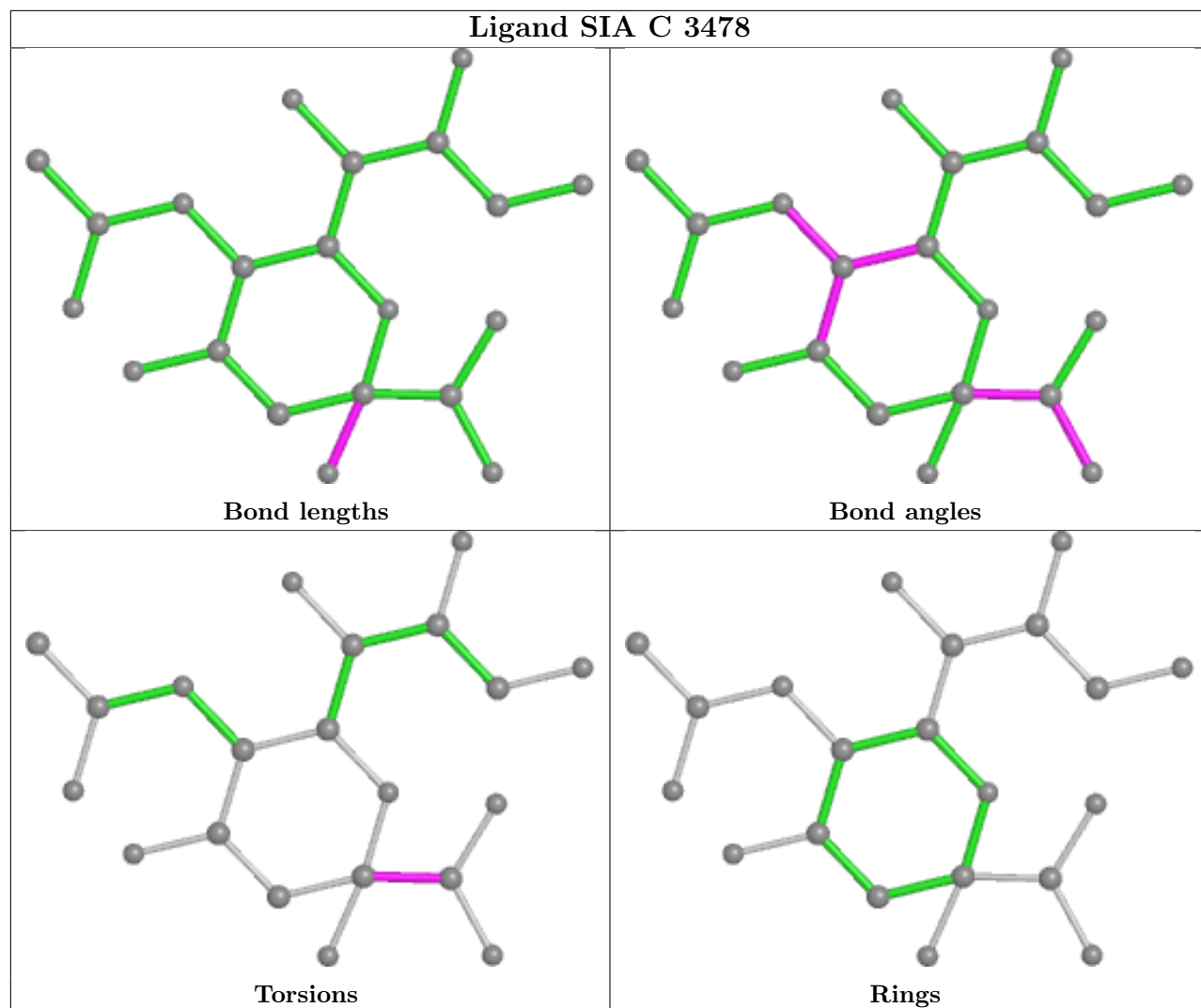
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In

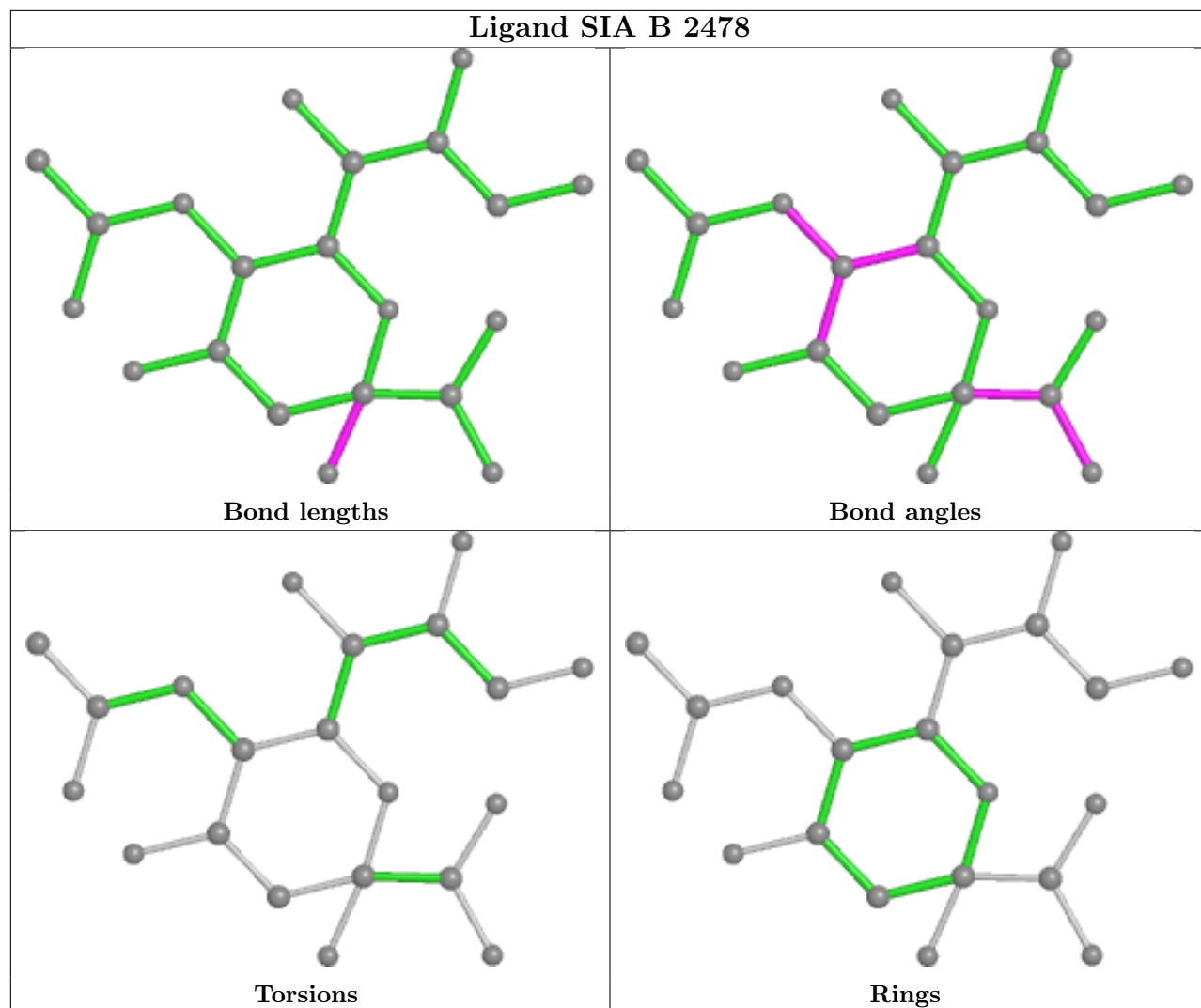
addition, ligands with molecular weight  $> 250$  and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. 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.

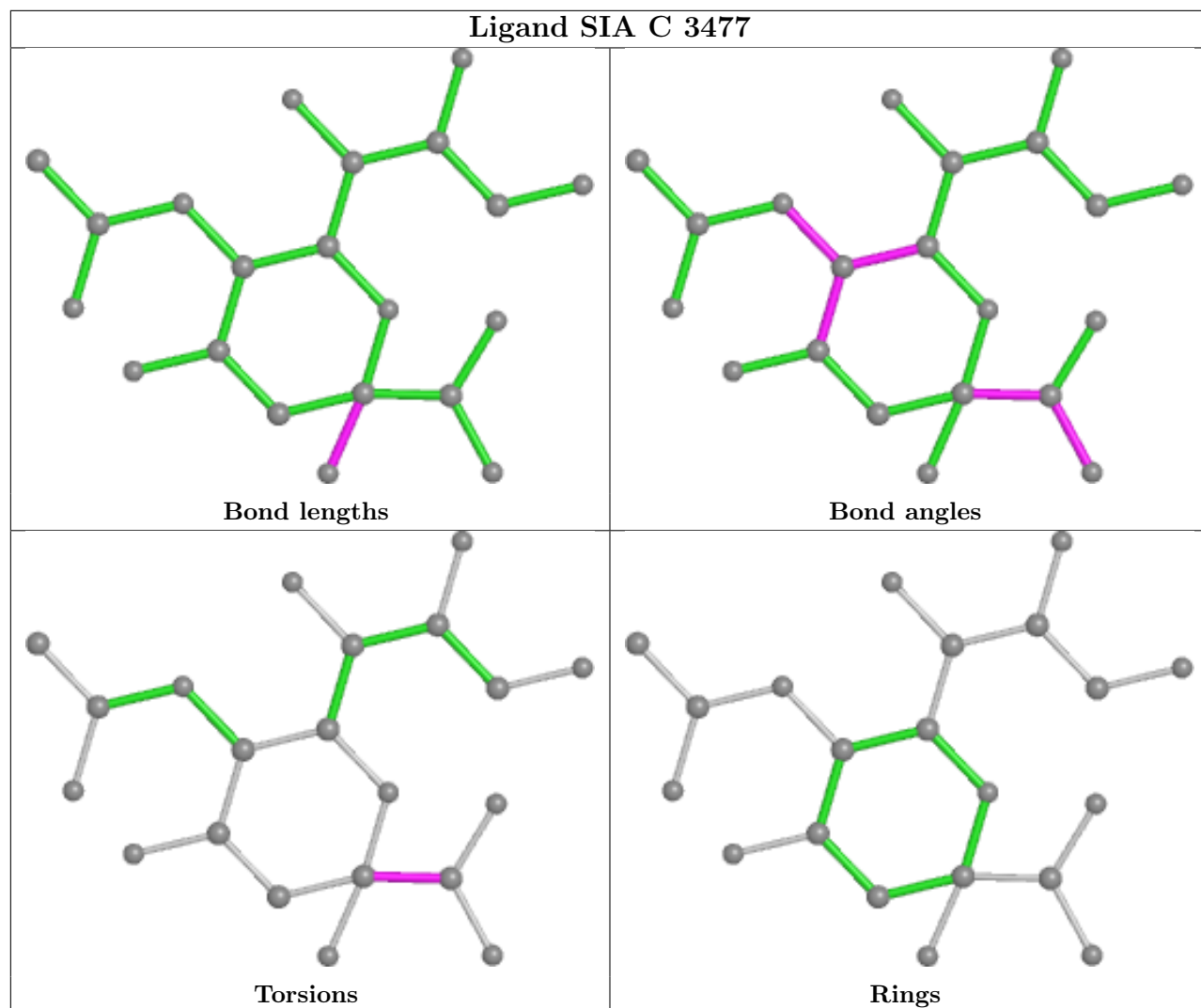


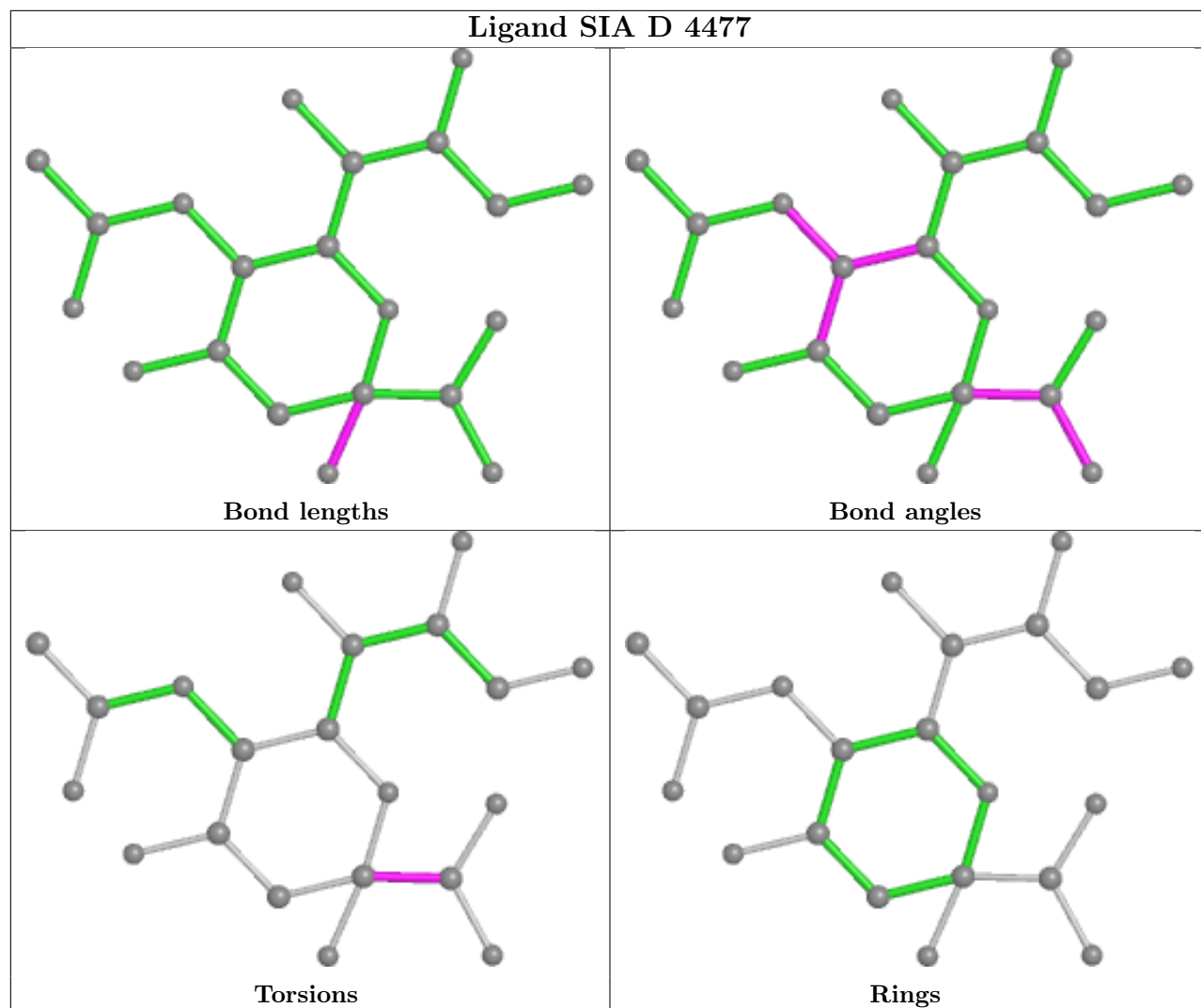




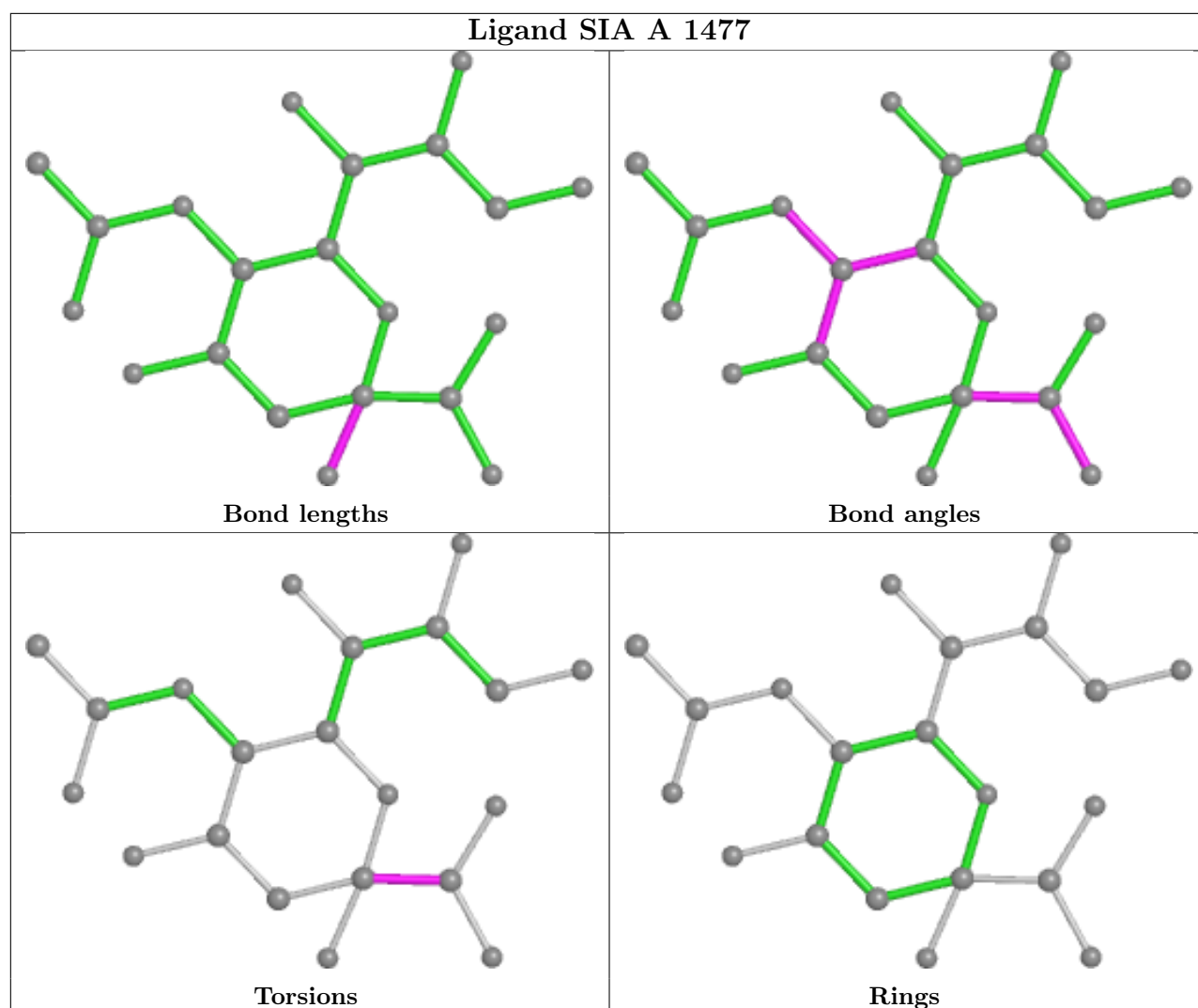












## 5.7 Other polymers [i](#)

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	389/389 (100%)	-0.39	1 (0%) 94   94	11, 16, 26, 41	0
1	B	389/389 (100%)	-0.37	3 (0%) 86   87	11, 17, 25, 48	0
1	C	389/389 (100%)	-0.44	2 (0%) 91   92	11, 17, 26, 41	0
1	D	389/389 (100%)	-0.38	1 (0%) 94   94	11, 17, 25, 38	0
All	All	1556/1556 (100%)	-0.40	7 (0%) 92   93	11, 17, 26, 48	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1089	THR	2.9
1	A	460	ARG	2.6
1	D	3460	ARG	2.4
1	B	1088	ARG	2.4
1	C	2088	ARG	2.3

### 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.

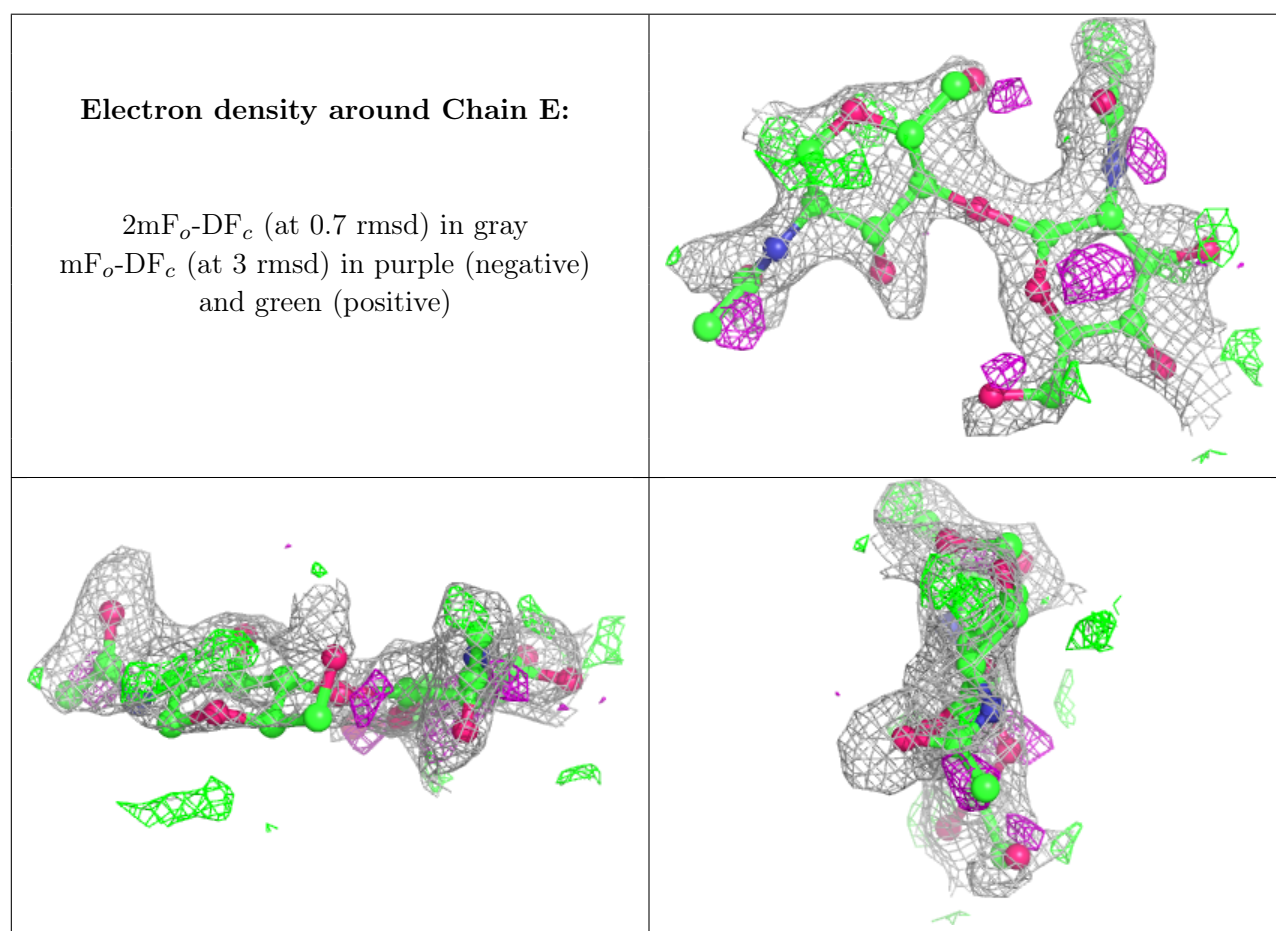
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAG	G	1	14/15	0.49	0.33	35,40,42,43	0
2	NAG	E	1	14/15	0.57	0.35	54,62,63,64	0

*Continued on next page...*

Continued from previous page...

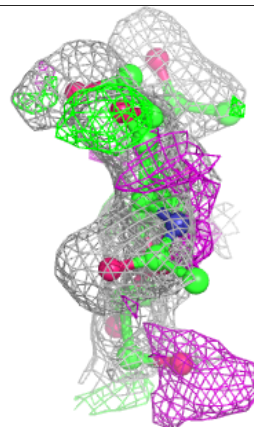
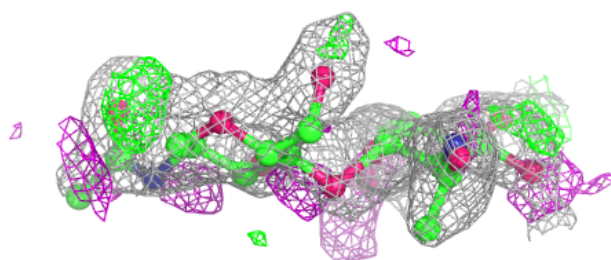
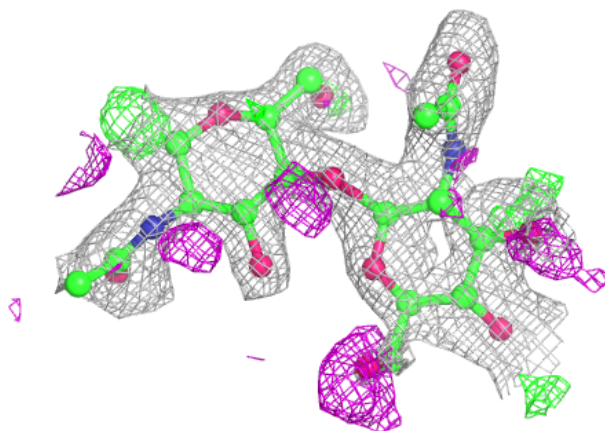
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	F	1	14/15	0.62	0.31	35,39,40,41	0
2	NAG	E	2	14/15	0.63	0.35	40,44,47,47	0
3	BMA	G	2	11/12	0.69	0.27	63,67,69,71	0
2	NAG	H	2	14/15	0.71	0.31	33,38,43,47	0
2	NAG	H	1	14/15	0.72	0.38	41,43,45,46	0
2	NAG	F	2	14/15	0.73	0.35	29,40,42,43	0
3	MAN	G	3	11/12	0.79	0.23	51,54,57,59	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.



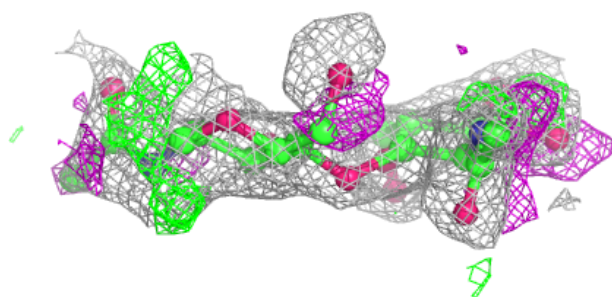
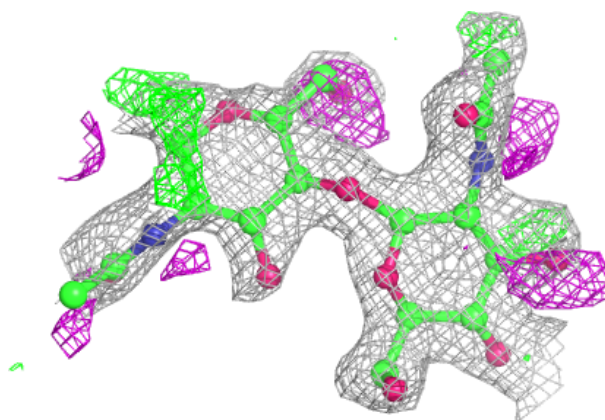
**Electron density around Chain F:**

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

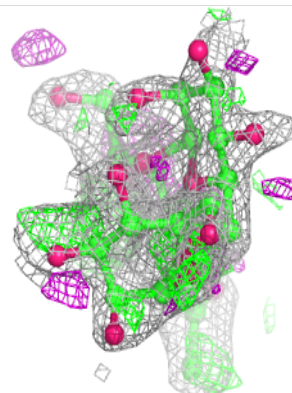
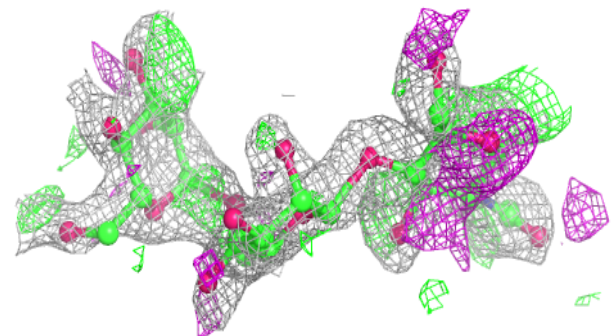
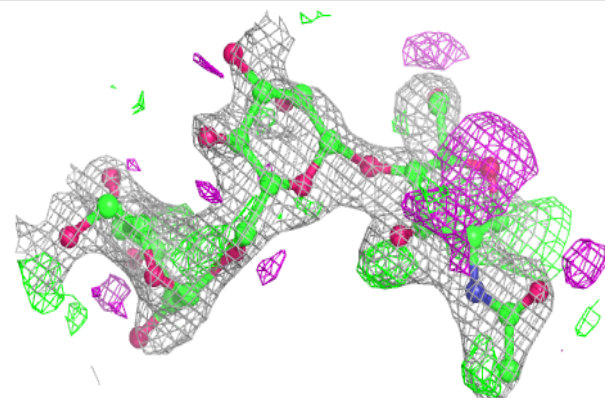


**Electron density around Chain H:**

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

**Electron density around Chain G:**

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



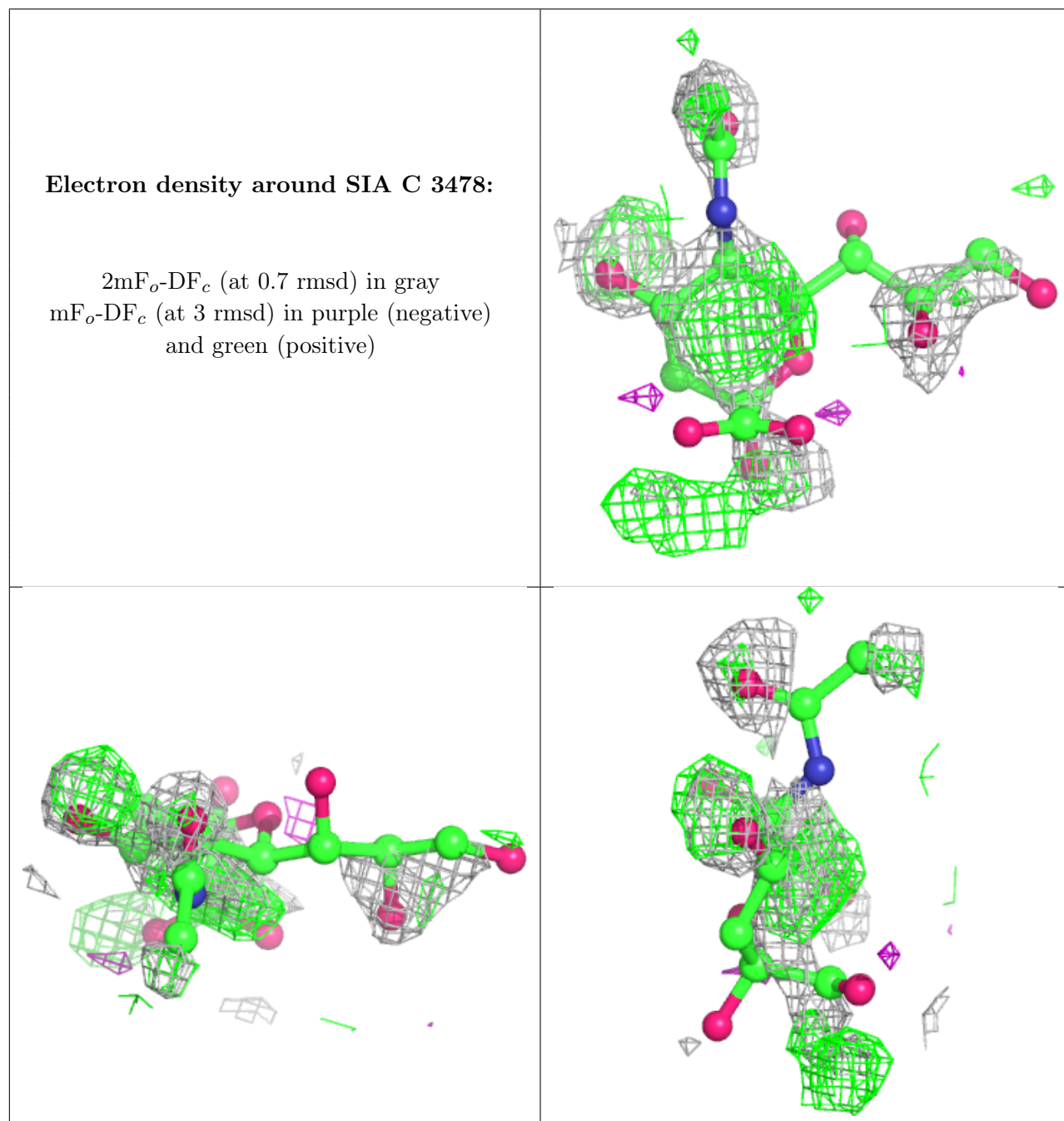
## 6.4 Ligands

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
8	MAN	C	3486	11/12	0.41	0.53	95,96,96,97	0
4	SIA	C	3478	21/21	0.47	0.72	28,32,34,36	21
7	BMA	A	1484	11/12	0.57	0.34	57,62,69,69	0
8	MAN	A	1486	11/12	0.64	0.27	55,59,64,66	0
8	MAN	B	2484	11/12	0.64	0.29	71,72,74,75	0
4	SIA	A	1478	21/21	0.64	0.46	16,19,22,23	21
5	GOL	C	3479	6/6	0.65	0.24	25,30,32,33	0
7	BMA	B	2483	11/12	0.69	0.40	61,68,72,72	0
5	GOL	B	2479	6/6	0.69	0.20	27,30,34,35	0
8	MAN	B	2486	11/12	0.72	0.35	67,73,75,76	0
8	MAN	A	1485	11/12	0.72	0.23	51,54,56,60	0
4	SIA	D	4478	21/21	0.73	0.23	22,33,35,39	21
8	MAN	C	3485	11/12	0.74	0.28	60,61,64,69	0
8	MAN	B	2485	11/12	0.74	0.30	62,64,66,67	0
8	MAN	A	1487	11/12	0.75	0.22	41,52,54,60	0
5	GOL	A	1479	6/6	0.75	0.21	20,32,33,33	0
7	BMA	C	3490	11/12	0.75	0.28	58,64,65,66	0
6	NAG	D	4480	14/15	0.77	0.22	25,33,37,38	0
5	GOL	D	4479	6/6	0.78	0.21	24,34,36,37	0
4	SIA	B	2478	21/21	0.79	0.21	24,32,37,37	21
8	MAN	C	3488	11/12	0.79	0.21	44,51,55,57	0
6	NAG	A	1480	14/15	0.80	0.23	36,42,44,45	0
6	NAG	C	3481	14/15	0.81	0.18	27,30,34,37	0
8	MAN	C	3489	11/12	0.81	0.20	39,41,47,51	0
4	SIA	A	1477	21/21	0.84	0.16	17,19,23,24	0
6	NAG	B	2480	14/15	0.85	0.20	21,33,39,43	0
4	SIA	C	3477	21/21	0.85	0.17	16,19,22,24	0
4	SIA	B	2477	21/21	0.87	0.15	17,19,21,24	0
4	SIA	D	4477	21/21	0.88	0.15	17,20,23,27	0
6	NAG	A	1481	14/15	0.88	0.23	24,32,38,42	0
9	CA	A	1488	1/1	0.98	0.27	8,8,8,8	0
9	CA	B	2487	1/1	0.98	0.32	5,5,5,5	0
9	CA	C	3491	1/1	0.98	0.34	2,2,2,2	0
9	CA	D	4481	1/1	0.99	0.28	4,4,4,4	0

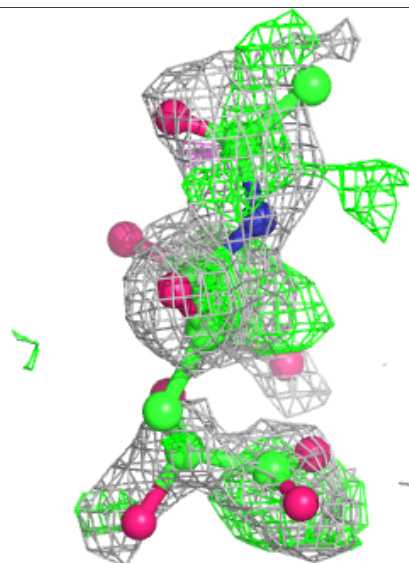
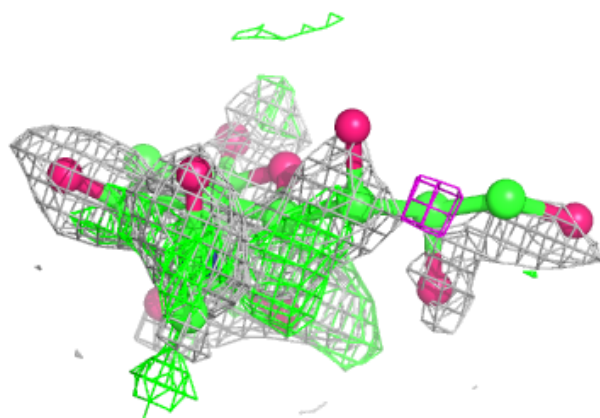
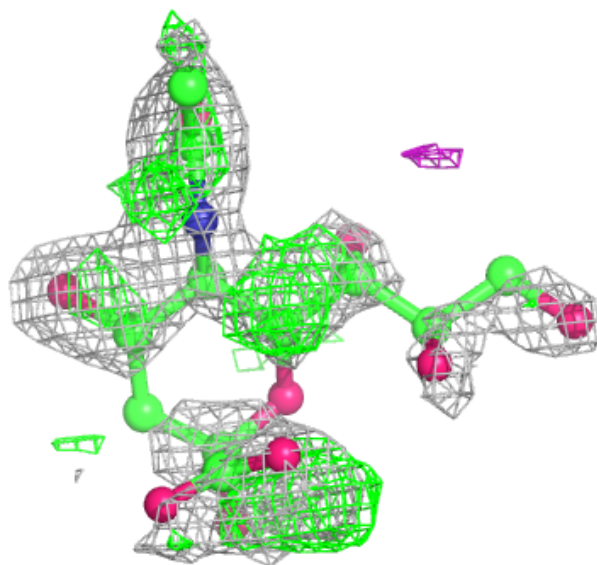
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

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 SIA A 1478:**

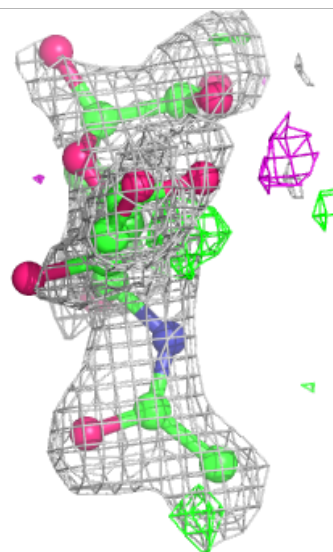
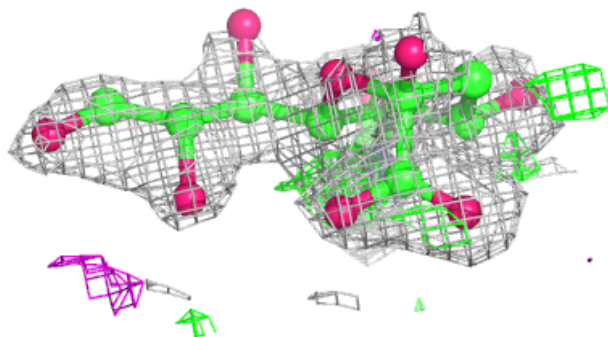
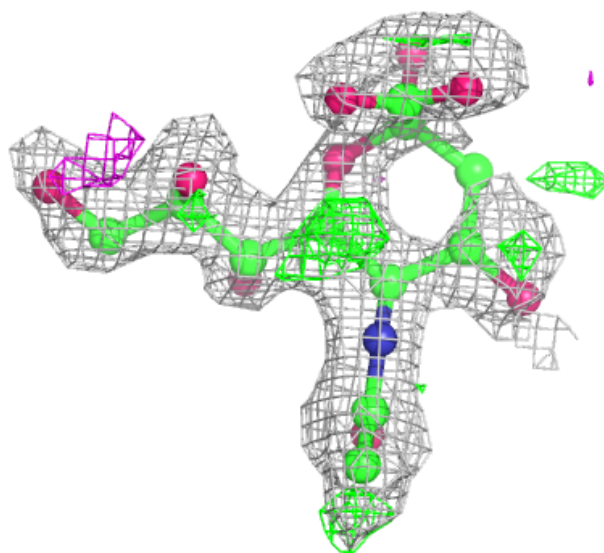
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





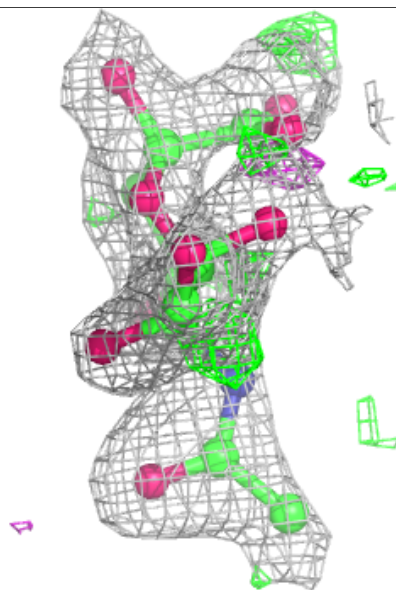
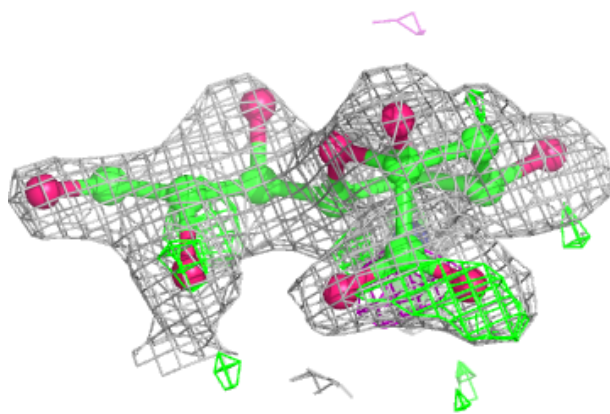
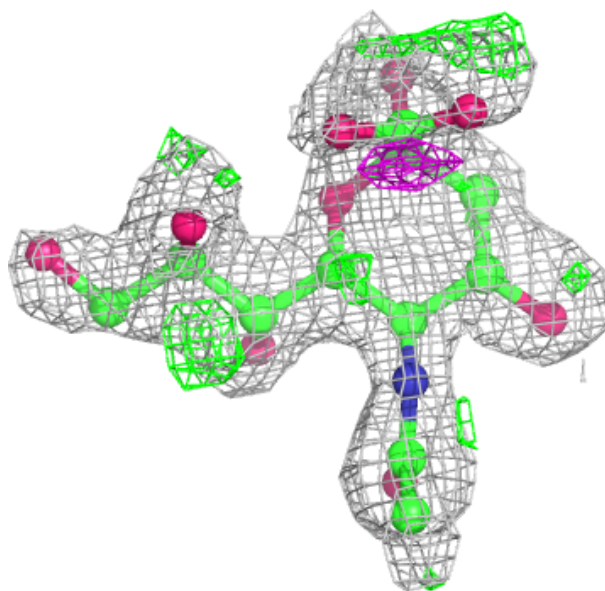
**Electron density around SIA D 4478:**

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



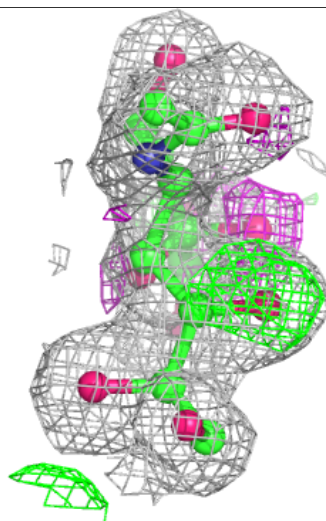
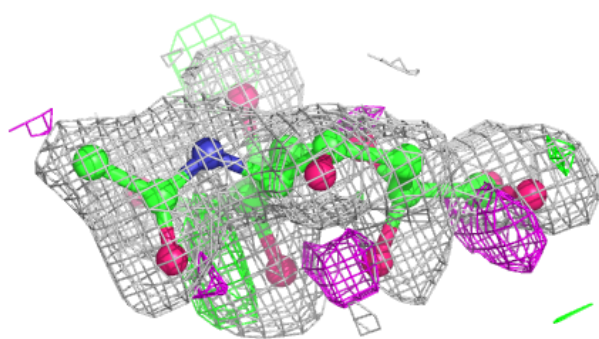
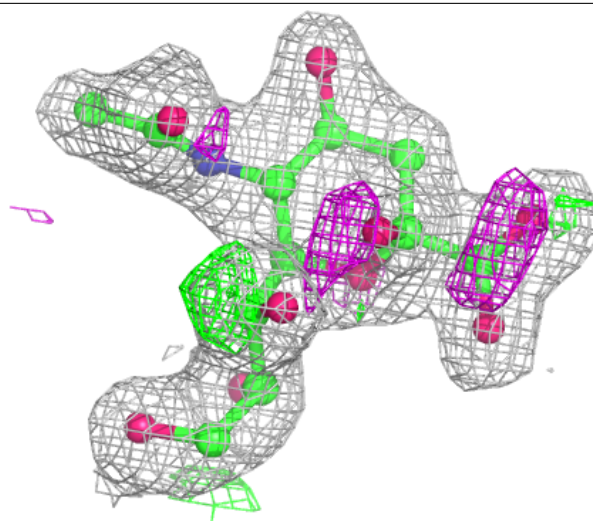
**Electron density around SIA B 2478:**

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



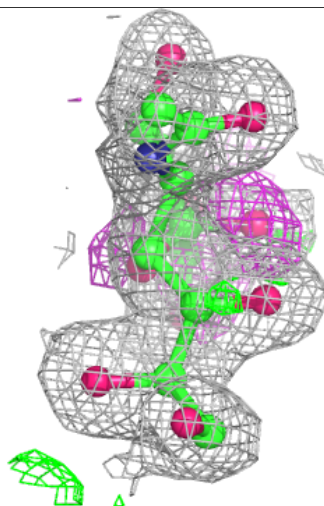
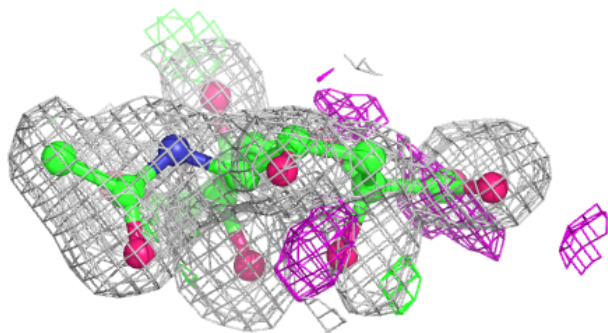
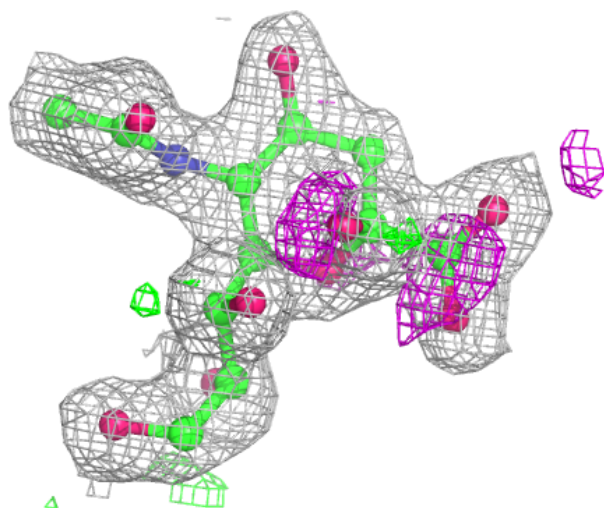
**Electron density around SIA A 1477:**

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



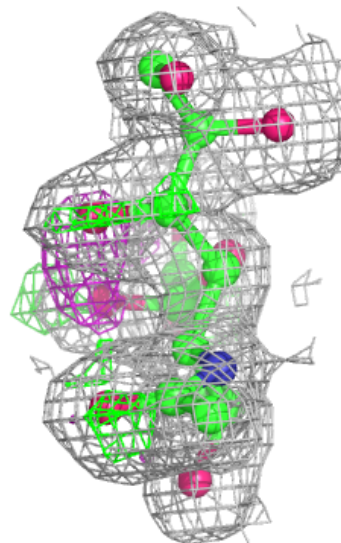
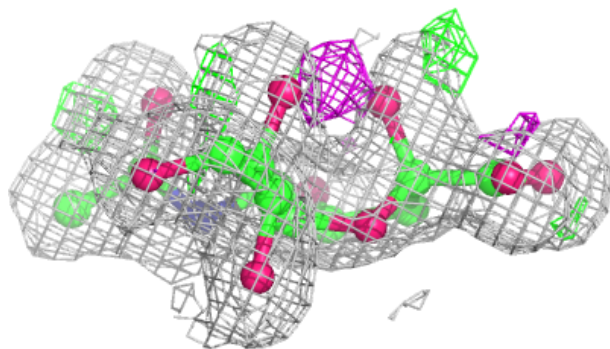
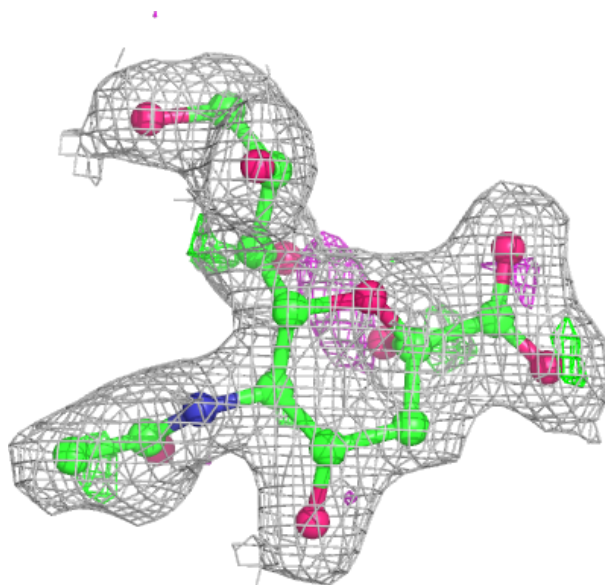
**Electron density around SIA C 3477:**

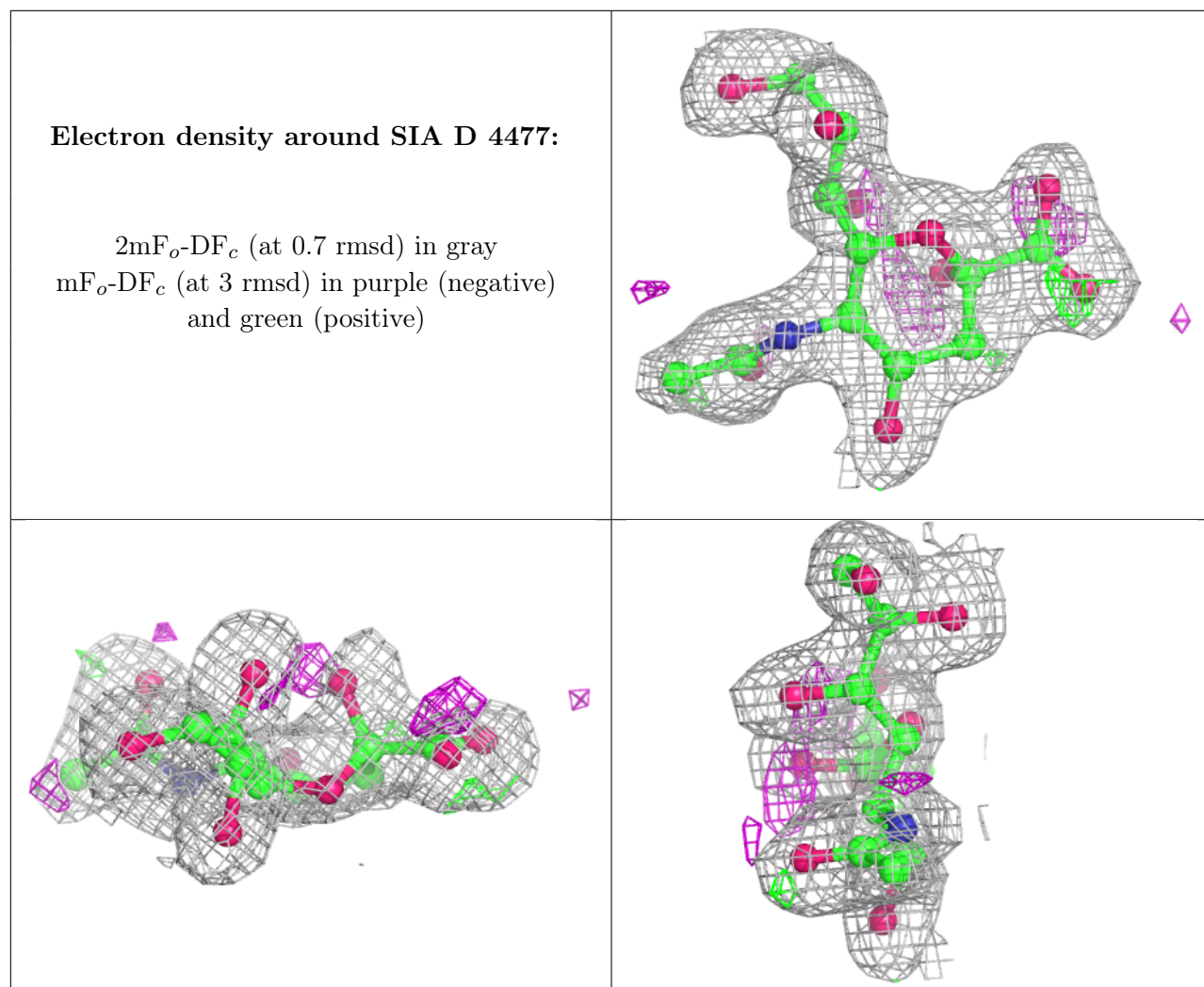
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around SIA B 2477:**

$2mF_o-DF_c$  (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.