



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 4, 2023 – 01:59 PM EDT

PDB ID : 7FQI  
Title : Crystal Structure of human Legumain in complex with (2S)-N-[(1S)-3-amino-1-cyano-3-oxopropyl]-1-[1-[4-[(2,4-difluorophenyl)methoxy]phenyl]cyclopropyl]pyrrolidine-2-carboxamide  
Authors : Ehler, A.; Benz, J.; Bartels, B.; Rudolph, M.G.  
Deposited on : 2022-10-05  
Resolution : 1.45 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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

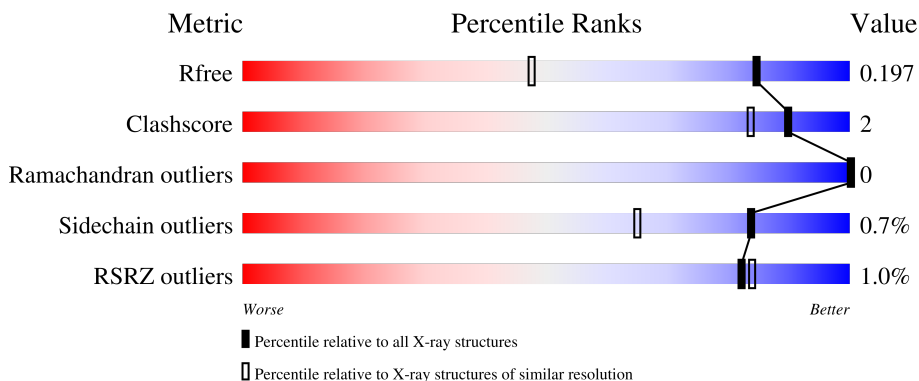
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.45 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.



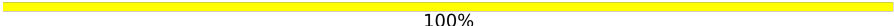

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1156 (1.46-1.46)
Clashscore	141614	1202 (1.46-1.46)
Ramachandran outliers	138981	1178 (1.46-1.46)
Sidechain outliers	138945	1178 (1.46-1.46)
RSRZ outliers	127900	1139 (1.46-1.46)

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	444	
1	B	444	
1	C	444	
2	D	2	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	E	2	 100%
2	F	2	 50% 50%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7347 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 Legumain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	265	2174	1373	373	413	15	0	7	0
1	B	264	2176	1380	367	414	15	0	8	0
1	C	261	2142	1361	364	402	15	0	7	0

There are 90 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q99538
A	1	LYS	-	expression tag	UNP Q99538
A	2	LEU	-	expression tag	UNP Q99538
A	3	CYS	-	expression tag	UNP Q99538
A	4	ILE	-	expression tag	UNP Q99538
A	5	LEU	-	expression tag	UNP Q99538
A	6	LEU	-	expression tag	UNP Q99538
A	7	ALA	-	expression tag	UNP Q99538
A	8	VAL	-	expression tag	UNP Q99538
A	9	VAL	-	expression tag	UNP Q99538
A	10	ALA	-	expression tag	UNP Q99538
A	11	PHE	-	expression tag	UNP Q99538
A	12	VAL	-	expression tag	UNP Q99538
A	13	GLY	-	expression tag	UNP Q99538
A	14	LEU	-	expression tag	UNP Q99538
A	15	SER	-	expression tag	UNP Q99538
A	16	LEU	-	expression tag	UNP Q99538
A	17	GLY	-	expression tag	UNP Q99538
A	147	SNN	ASP	conflict	UNP Q99538
A	272	GLN	ASN	conflict	UNP Q99538
A	434	VAL	-	expression tag	UNP Q99538
A	435	ASP	-	expression tag	UNP Q99538
A	436	HIS	-	expression tag	UNP Q99538

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	437	HIS	-	expression tag	UNP Q99538
A	438	HIS	-	expression tag	UNP Q99538
A	439	HIS	-	expression tag	UNP Q99538
A	440	HIS	-	expression tag	UNP Q99538
A	441	HIS	-	expression tag	UNP Q99538
A	442	HIS	-	expression tag	UNP Q99538
A	443	HIS	-	expression tag	UNP Q99538
B	0	MET	-	initiating methionine	UNP Q99538
B	1	LYS	-	expression tag	UNP Q99538
B	2	LEU	-	expression tag	UNP Q99538
B	3	CYS	-	expression tag	UNP Q99538
B	4	ILE	-	expression tag	UNP Q99538
B	5	LEU	-	expression tag	UNP Q99538
B	6	LEU	-	expression tag	UNP Q99538
B	7	ALA	-	expression tag	UNP Q99538
B	8	VAL	-	expression tag	UNP Q99538
B	9	VAL	-	expression tag	UNP Q99538
B	10	ALA	-	expression tag	UNP Q99538
B	11	PHE	-	expression tag	UNP Q99538
B	12	VAL	-	expression tag	UNP Q99538
B	13	GLY	-	expression tag	UNP Q99538
B	14	LEU	-	expression tag	UNP Q99538
B	15	SER	-	expression tag	UNP Q99538
B	16	LEU	-	expression tag	UNP Q99538
B	17	GLY	-	expression tag	UNP Q99538
B	147	SNN	ASP	conflict	UNP Q99538
B	272	GLN	ASN	conflict	UNP Q99538
B	434	VAL	-	expression tag	UNP Q99538
B	435	ASP	-	expression tag	UNP Q99538
B	436	HIS	-	expression tag	UNP Q99538
B	437	HIS	-	expression tag	UNP Q99538
B	438	HIS	-	expression tag	UNP Q99538
B	439	HIS	-	expression tag	UNP Q99538
B	440	HIS	-	expression tag	UNP Q99538
B	441	HIS	-	expression tag	UNP Q99538
B	442	HIS	-	expression tag	UNP Q99538
B	443	HIS	-	expression tag	UNP Q99538
C	0	MET	-	initiating methionine	UNP Q99538
C	1	LYS	-	expression tag	UNP Q99538
C	2	LEU	-	expression tag	UNP Q99538
C	3	CYS	-	expression tag	UNP Q99538
C	4	ILE	-	expression tag	UNP Q99538

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	5	LEU	-	expression tag	UNP Q99538
C	6	LEU	-	expression tag	UNP Q99538
C	7	ALA	-	expression tag	UNP Q99538
C	8	VAL	-	expression tag	UNP Q99538
C	9	VAL	-	expression tag	UNP Q99538
C	10	ALA	-	expression tag	UNP Q99538
C	11	PHE	-	expression tag	UNP Q99538
C	12	VAL	-	expression tag	UNP Q99538
C	13	GLY	-	expression tag	UNP Q99538
C	14	LEU	-	expression tag	UNP Q99538
C	15	SER	-	expression tag	UNP Q99538
C	16	LEU	-	expression tag	UNP Q99538
C	17	GLY	-	expression tag	UNP Q99538
C	147	SNN	ASP	conflict	UNP Q99538
C	272	GLN	ASN	conflict	UNP Q99538
C	434	VAL	-	expression tag	UNP Q99538
C	435	ASP	-	expression tag	UNP Q99538
C	436	HIS	-	expression tag	UNP Q99538
C	437	HIS	-	expression tag	UNP Q99538
C	438	HIS	-	expression tag	UNP Q99538
C	439	HIS	-	expression tag	UNP Q99538
C	440	HIS	-	expression tag	UNP Q99538
C	441	HIS	-	expression tag	UNP Q99538
C	442	HIS	-	expression tag	UNP Q99538
C	443	HIS	-	expression tag	UNP Q99538

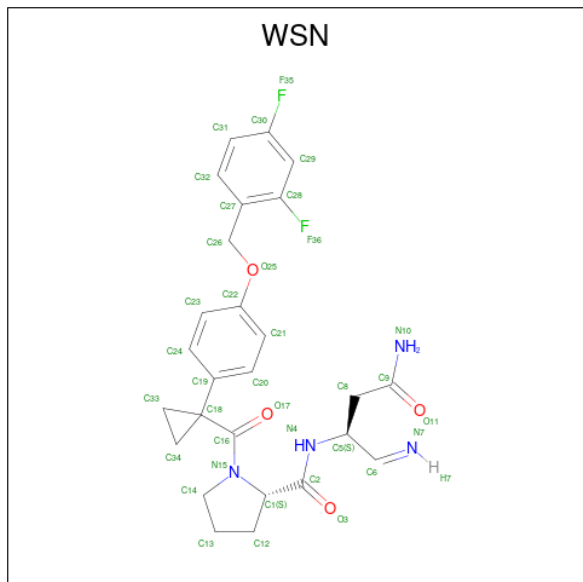
- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



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

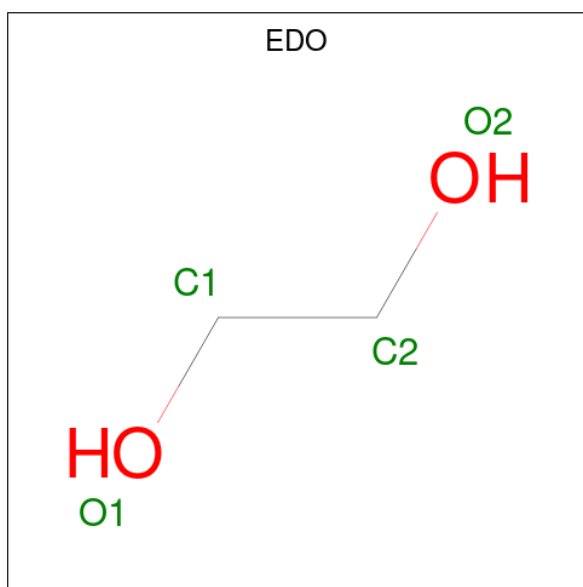
- Molecule 3 is N-[(2R)-4-amino-1-imino-4-oxobutan-2-yl]-1-(1-{4-[(2,4-difluorophenyl)met

hoxy]phenyl}cyclopropane-1-carbonyl)-L-prolinamide (three-letter code: WSN) (formula:  $C_{26}H_{28}F_2N_4O_4$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf		
			Total	C	F	N			O	
3	A	1	Total	36	26	2	4	4	0	0
3	B	1	Total	45	33	4	4	4	0	1
3	C	1	Total	36	26	2	4	4	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	237	Total O 239 239	0	2
5	B	213	Total O 214 214	0	1
5	C	181	Total O 181 181	0	0





SER PRO VAL SER LEU PRO PRO VAL THR THR LEU ASP LEU THR PRO SER PRO ASP VAL PRO LEU THR THR ILE MET LYS ARG LYS LEU MET ASN THR ASN ASP LEU LEU PHE GLU GLU SER ARG GLN LEU THR GLU THR ARG HIS LEU LEU TRP TRP ASN THR HIS LEU SER ILE GLN PRO ARG TYR HIS LEU ASP ALA ARG LEU HIS LEU ILE LEU TYR VAL LYS SER CYS TYR PRO GLU ALA LEU LEU HIS PHE ARG THR HIS CYS PHE LEU ASN HIS TRP SER PRO TYR TYR ALA LEU ARG HIS LEU ILE LEU TYR VAL LYS SER VAL CYS LEU GLY HIS TYR VAL ASP HIS HIS HIS HIS HIS HIS HIS HIS HIS HIS

ILE VAL SER LEU ALA ALA SER ARG ILE GLU VAL GLN LEU SER LEU ARG ALA PRO LEU THR ILE THR GLY MET LYS SER CYS TYR PRO GLU ALA LEU LEU HIS PHE ARG THR HIS CYS PHE LEU ASN HIS TRP SER PRO TYR TYR ALA LEU ARG HIS LEU ILE LEU TYR VAL LYS SER VAL CYS LEU GLY HIS TYR VAL ASP HIS HIS HIS HIS HIS HIS HIS HIS HIS HIS

CYS  
GLU  
LYS  
PRO  
TYR  
PRO  
LEU  
HIS  
ARG  
ILE  
LYS  
LEU  
SER  
MET  
ASP  
HIS  
VAL  
CYS  
LEU  
GLY  
HIS  
TYR  
VAL  
ASP  
HIS  
HIS  
HIS  
HIS  
HIS  
HIS  
HIS  
HIS

- 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



MAG1  
MAG2

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



MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	229.51Å 67.09Å 79.12Å 90.00° 100.63° 90.00°	Depositor
Resolution (Å)	70.37 – 1.45 70.37 – 1.45	Depositor EDS
% Data completeness (in resolution range)	67.0 (70.37-1.45) 67.0 (70.37-1.45)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.68 (at 1.45Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.167 , 0.189 0.176 , 0.197	Depositor DCC
$R_{free}$ test set	7335 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.9	Xtrriage
Anisotropy	0.024	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 39.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7347	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.79% 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: SNN, WSN, EDO, 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	0.66	0/2241	0.79	0/3035
1	B	0.65	0/2247	0.78	0/3043
1	C	0.65	0/2212	0.76	0/2998
All	All	0.65	0/6700	0.78	0/9076

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	2174	0	2073	6	0
1	B	2176	0	2076	10	0
1	C	2142	0	2061	7	0
2	D	28	0	25	0	0
2	E	28	0	25	0	0
2	F	28	0	25	0	0
3	A	36	0	0	0	0
3	B	45	0	0	0	0
3	C	36	0	0	0	0
4	A	12	0	18	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	8	0	12	1	0
5	A	239	0	0	2	0
5	B	214	0	0	3	0
5	C	181	0	0	2	0
All	All	7347	0	6315	23	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.

The worst 5 of 23 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:282[B]:GLN:OE1	5:A:601:HOH:O	1.77	1.01
1:A:89[A]:ARG:HD2	5:A:806:HOH:O	1.81	0.81
1:B:122[A]:GLU:OE1	5:B:602:HOH:O	2.03	0.77
1:C:253[A]:LYS:NZ	5:C:601:HOH:O	2.02	0.60
1:B:41[A]:TYR:OH	5:B:601:HOH:O	0.60	0.59

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	269/444 (61%)	263 (98%)	6 (2%)	0	100	100
1	B	269/444 (61%)	261 (97%)	8 (3%)	0	100	100
1	C	265/444 (60%)	260 (98%)	5 (2%)	0	100	100
All	All	803/1332 (60%)	784 (98%)	19 (2%)	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	237/396 (60%)	235 (99%)	2 (1%)	81	62
1	B	238/396 (60%)	236 (99%)	2 (1%)	81	62
1	C	235/396 (59%)	233 (99%)	2 (1%)	78	57
All	All	710/1188 (60%)	704 (99%)	6 (1%)	84	62

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	263	ASN
1	C	97[A]	GLN
1	C	97[B]	GLN
1	A	134	LYS
1	A	101	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	197	HIS
1	B	263	ASN
1	C	88	ASN
1	B	65	GLN
1	A	88	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

3 non-standard protein/DNA/RNA residues 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$
1	SNN	C	147	1	5,6,8	1.02	0	3,6,11	3.30	2 (66%)
1	SNN	B	147	1	5,6,8	0.93	0	3,6,11	3.42	2 (66%)
1	SNN	A	147	1	5,6,8	0.88	0	3,6,11	3.04	2 (66%)

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
1	SNN	C	147	1	-	3/3/5/12	-
1	SNN	B	147	1	-	3/3/5/12	-
1	SNN	A	147	1	-	3/3/5/12	-

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	147	SNN	C4-CA-C	-4.96	102.18	111.44
1	B	147	SNN	C4-CA-C	-4.56	102.93	111.44
1	A	147	SNN	CA-C4-C5	-3.77	100.16	114.44
1	B	147	SNN	CA-C4-C5	-3.66	100.59	114.44
1	A	147	SNN	C4-CA-C	-3.54	104.83	111.44

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	147	SNN	O-C-CA-C4
1	A	147	SNN	C5-C4-CA-N
1	B	147	SNN	O-C-CA-C4
1	B	147	SNN	C5-C4-CA-N
1	C	147	SNN	O-C-CA-C4

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

6 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	D	1	1,2	14,14,15	0.49	0	17,19,21	0.62	0
2	NAG	D	2	2	14,14,15	0.43	0	17,19,21	1.64	2 (11%)
2	NAG	E	1	1,2	14,14,15	0.61	0	17,19,21	1.09	2 (11%)
2	NAG	E	2	2	14,14,15	0.40	0	17,19,21	1.21	1 (5%)
2	NAG	F	1	1,2	14,14,15	0.47	0	17,19,21	0.83	0
2	NAG	F	2	2	14,14,15	0.35	0	17,19,21	1.33	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	D	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	D	2	2	-	0/6/23/26	0/1/1/1
2	NAG	E	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1
2	NAG	F	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	F	2	2	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	2	NAG	C1-O5-C5	5.41	119.53	112.19

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	2	NAG	C1-O5-C5	4.78	118.67	112.19
2	E	2	NAG	C1-O5-C5	3.12	116.42	112.19
2	D	2	NAG	O5-C1-C2	3.04	116.09	111.29
2	E	1	NAG	O5-C5-C6	2.40	110.97	107.20

There are no chirality outliers.

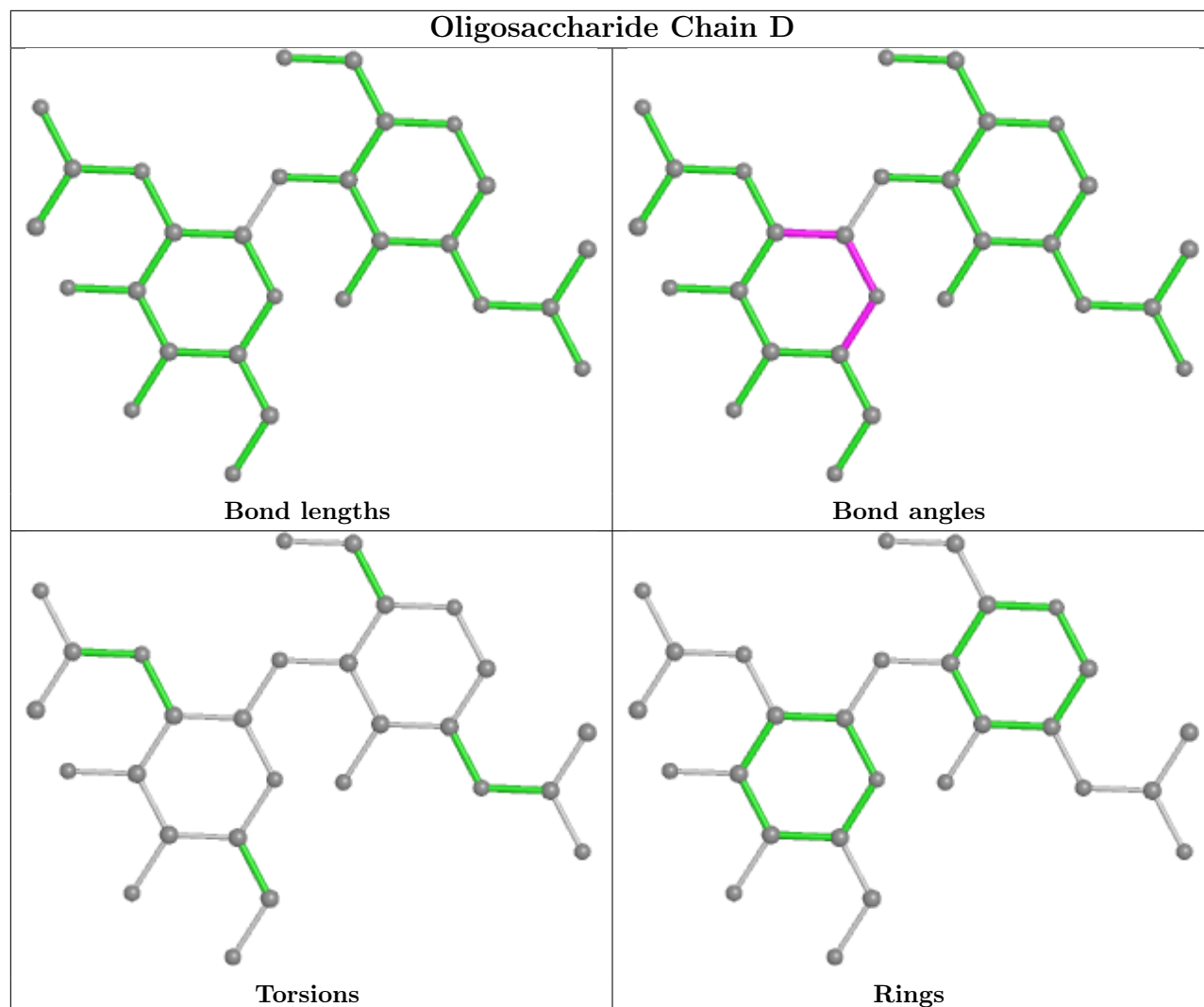
All (2) torsion outliers are listed below:

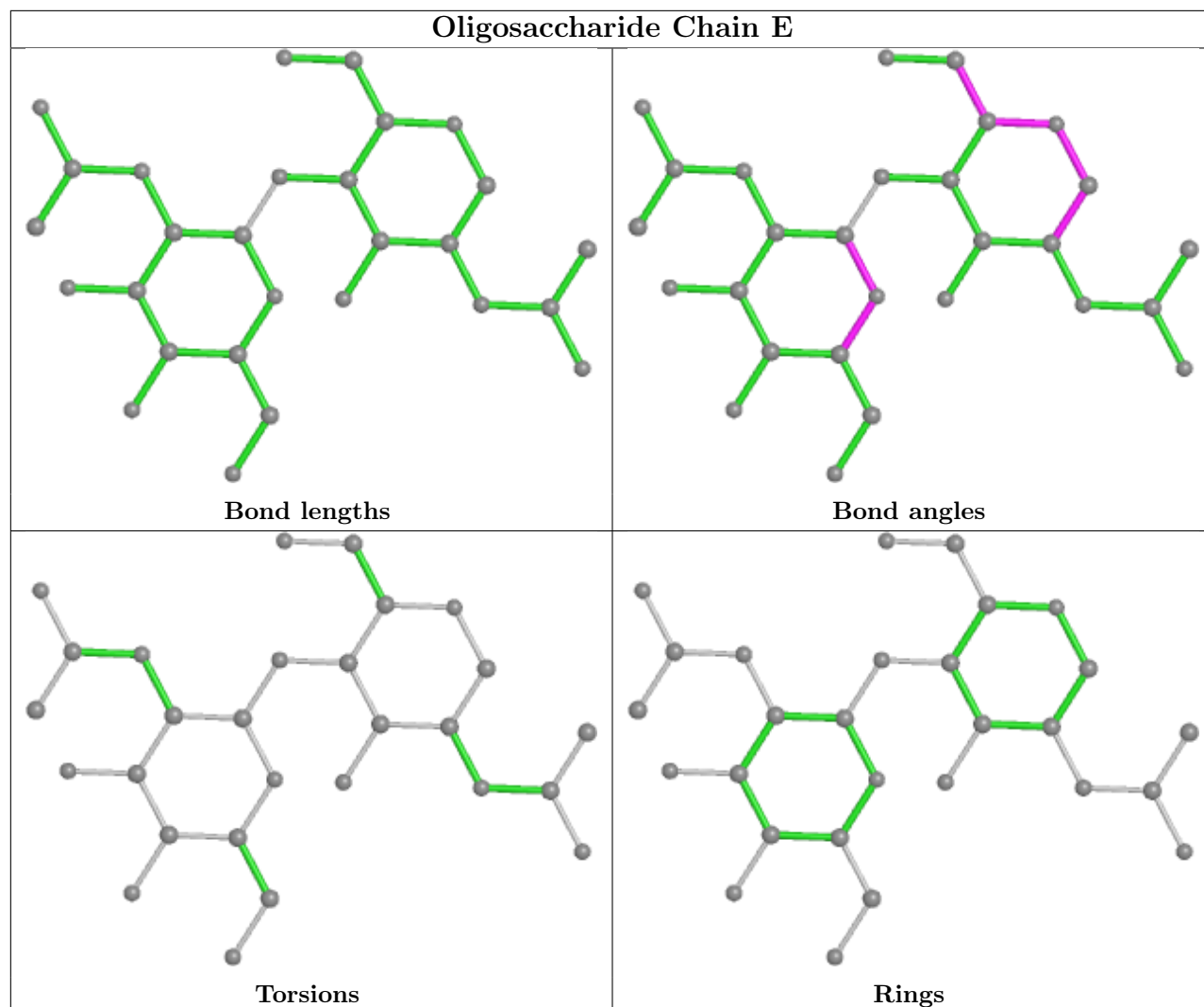
Mol	Chain	Res	Type	Atoms
2	F	2	NAG	C4-C5-C6-O6
2	F	2	NAG	O5-C5-C6-O6

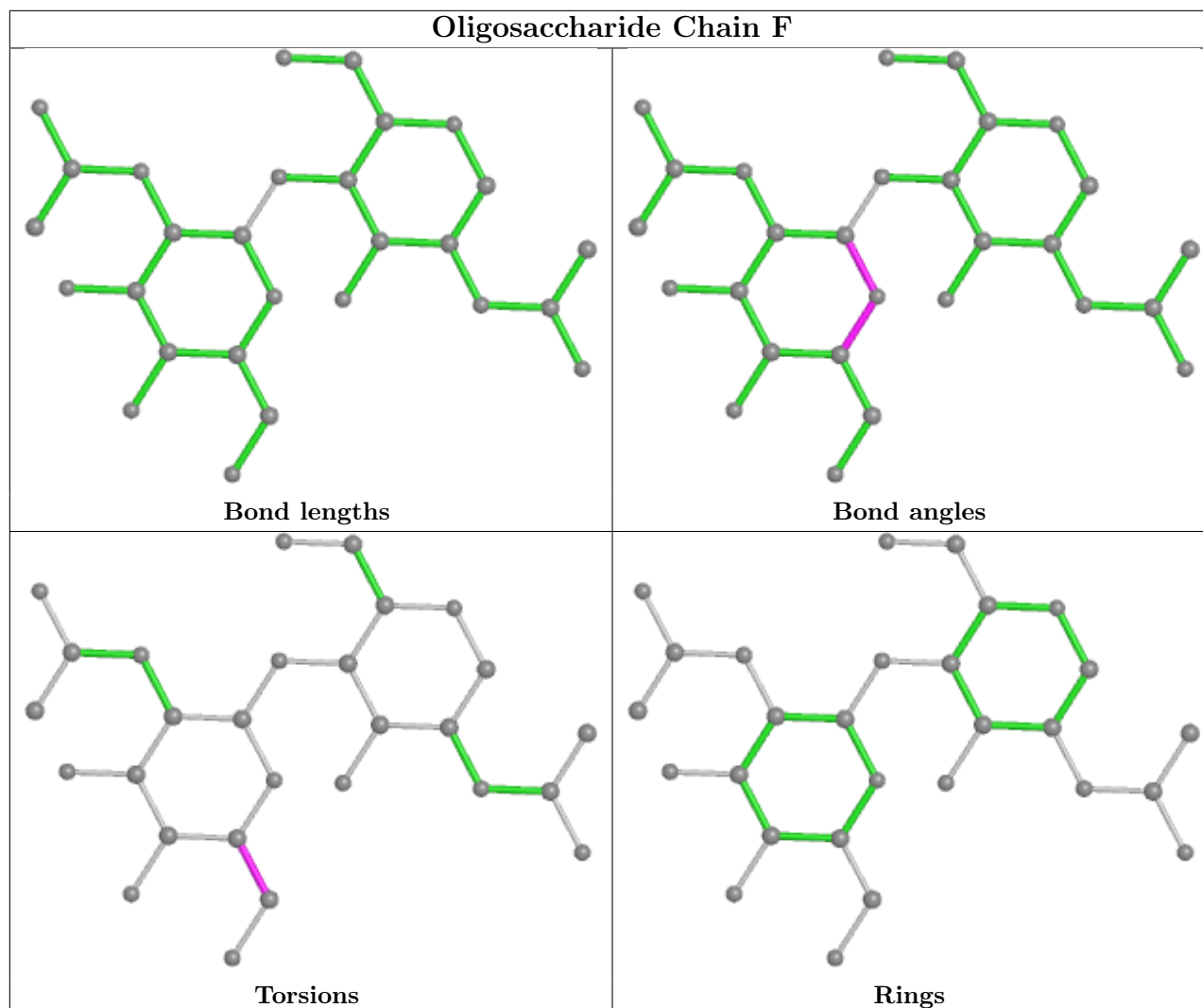
There are no ring outliers.

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 oligosaccharide.







## 5.6 Ligand geometry [i](#)

9 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	WSN	B	501[B]	-	36,39,39	0.30	0	49,56,56	0.65	1 (2%)
4	EDO	A	503	-	3,3,3	0.25	0	2,2,2	0.38	0
4	EDO	A	504	-	3,3,3	0.29	0	2,2,2	0.36	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	EDO	C	503	-	3,3,3	0.40	0	2,2,2	0.37	0
4	EDO	C	502	-	3,3,3	0.19	0	2,2,2	0.08	0
4	EDO	A	502	-	3,3,3	0.58	0	2,2,2	0.53	0
3	WSN	A	501	1	36,39,39	0.45	0	49,56,56	0.75	1 (2%)
3	WSN	B	501[A]	-	36,39,39	0.27	0	49,56,56	0.57	0
3	WSN	C	501	1	36,39,39	0.41	0	49,56,56	0.76	2 (4%)

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	WSN	B	501[B]	-	-	3/33/49/49	0/4/4/4
4	EDO	A	503	-	-	0/1/1/1	-
4	EDO	A	504	-	-	0/1/1/1	-
4	EDO	C	503	-	-	1/1/1/1	-
4	EDO	C	502	-	-	0/1/1/1	-
4	EDO	A	502	-	-	0/1/1/1	-
3	WSN	A	501	1	-	0/33/49/49	0/4/4/4
3	WSN	B	501[A]	-	-	0/33/49/49	0/4/4/4
3	WSN	C	501	1	-	3/33/49/49	0/4/4/4

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	501	WSN	O25-C26-C27	2.66	117.23	109.38
3	C	501	WSN	C26-C27-C32	2.42	124.78	119.52
3	B	501[B]	WSN	O25-C26-C27	2.08	115.51	109.38
3	C	501	WSN	C5-N4-C2	2.03	126.89	123.15

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	501[B]	WSN	C23-C22-O25-C26
3	B	501[B]	WSN	C21-C22-O25-C26
3	B	501[B]	WSN	O25-C26-C27-C28
3	C	501	WSN	C33-C18-C19-C24

*Continued on next page...*

*Continued from previous page...*

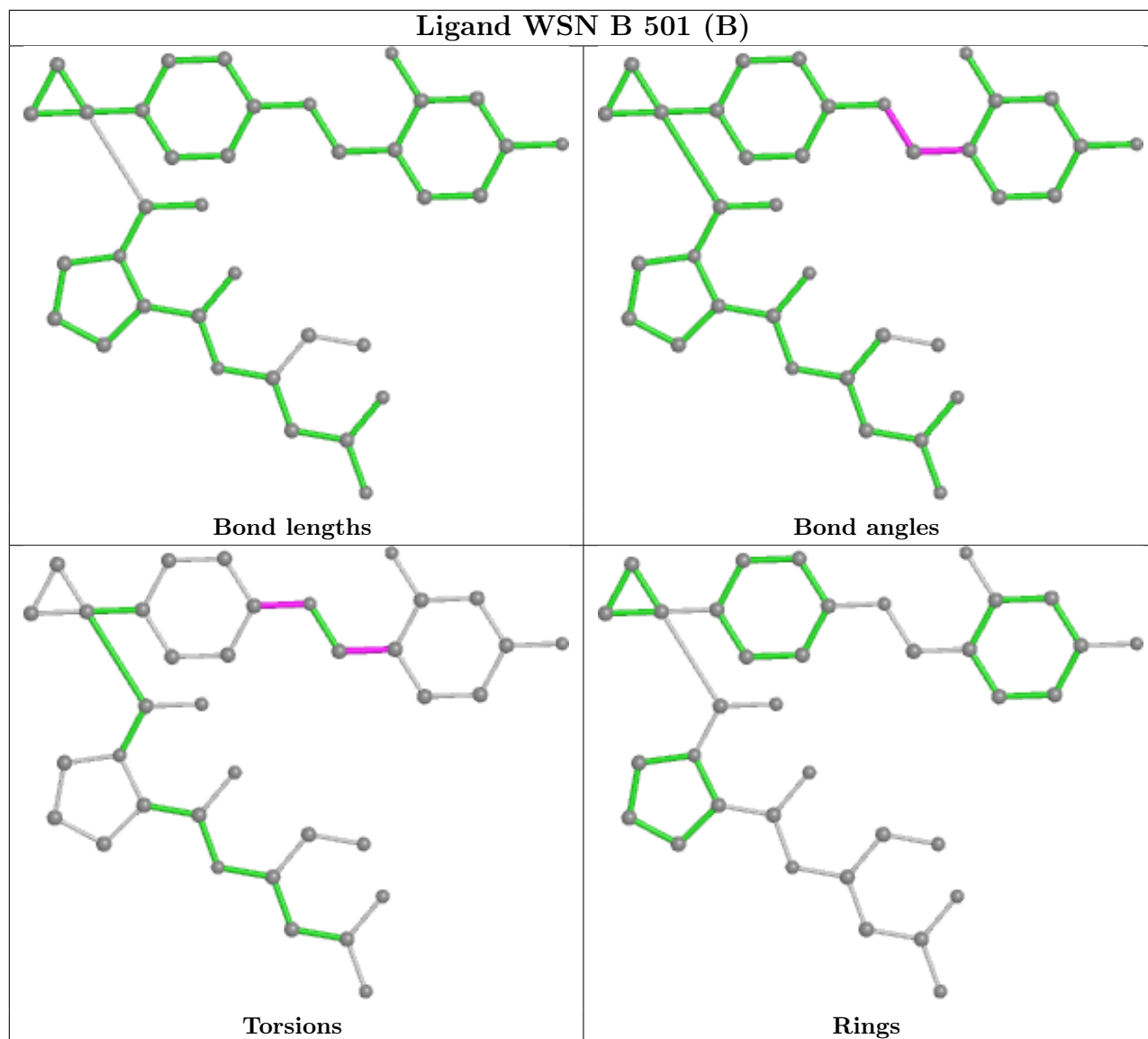
Mol	Chain	Res	Type	Atoms
3	C	501	WSN	C21-C22-O25-C26

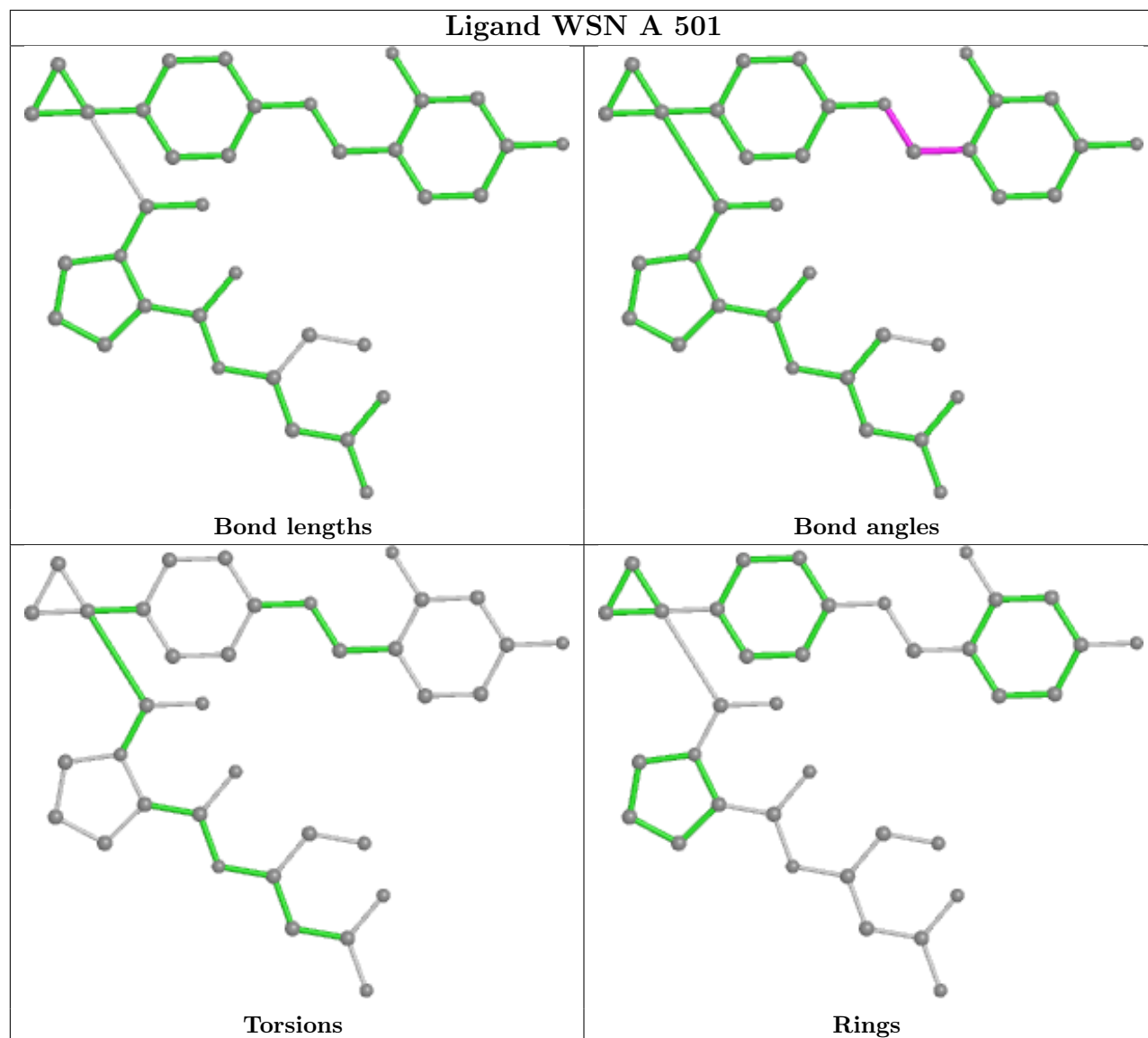
There are no ring outliers.

1 monomer is involved in 1 short contact:

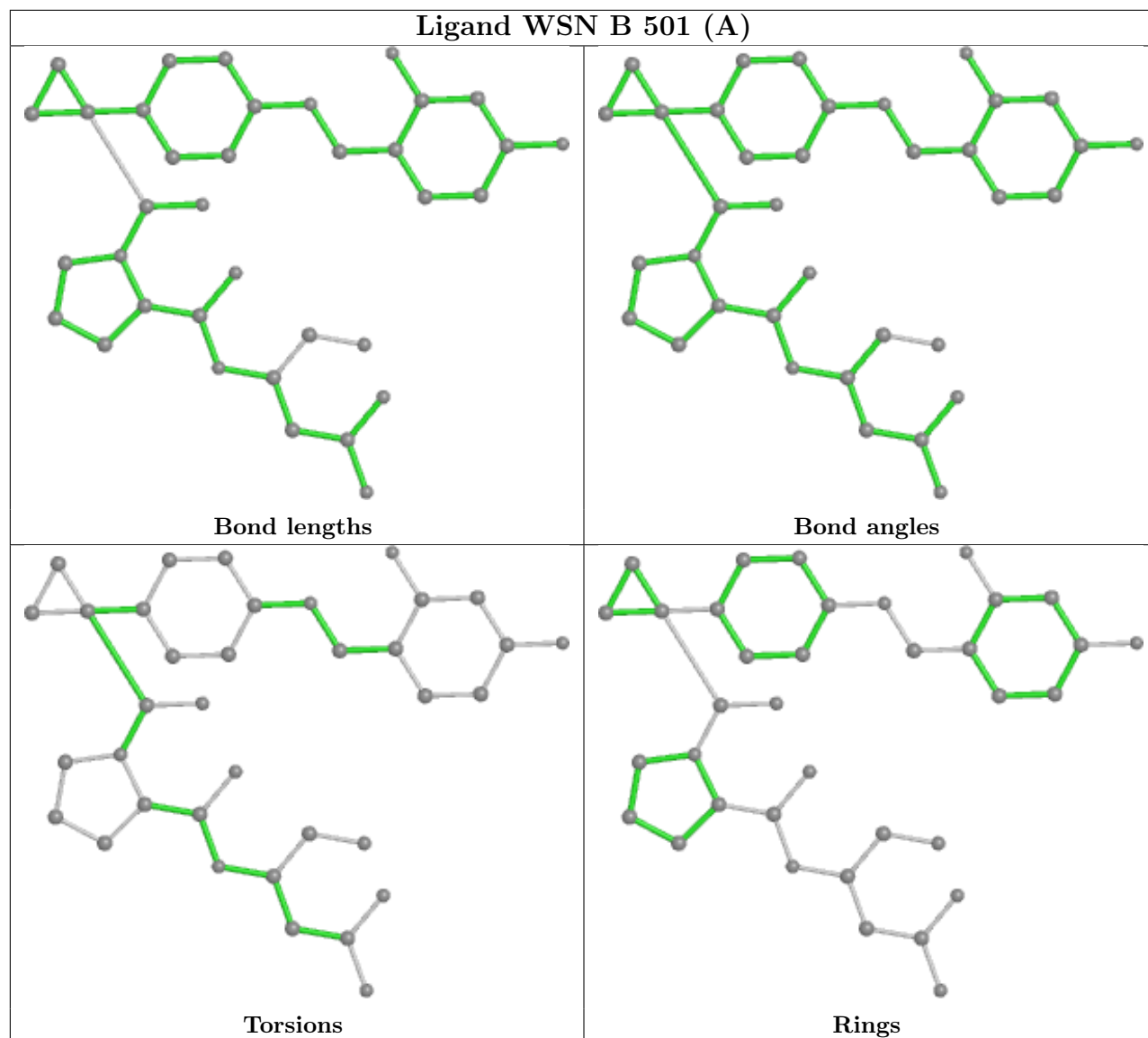
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	502	EDO	1	0

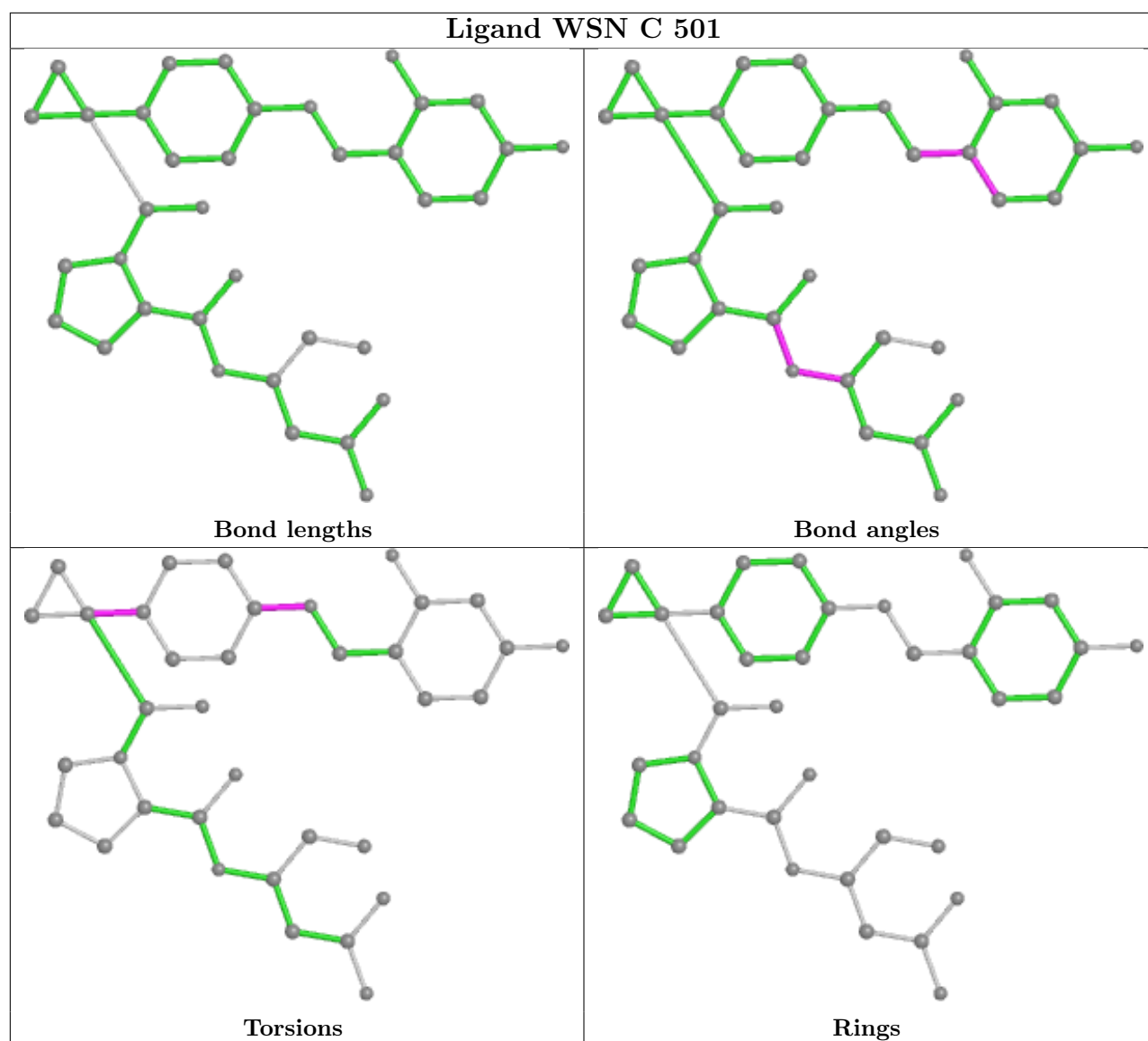
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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	264/444 (59%)	-0.23	3 (1%) 80 82	12, 18, 33, 53	0
1	B	263/444 (59%)	-0.27	4 (1%) 73 74	14, 20, 33, 72	0
1	C	260/444 (58%)	-0.15	1 (0%) 92 94	14, 22, 40, 59	0
All	All	787/1332 (59%)	-0.21	8 (1%) 82 84	12, 20, 34, 72	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	290	ALA	6.1
1	B	289	LYS	5.1
1	A	26	GLY	2.8
1	B	41[A]	TYR	2.7
1	B	26	GLY	2.6

### 6.2 Non-standard residues in protein, DNA, RNA chains [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
1	SNN	A	147	7/8	0.97	0.08	11,12,14,14	0
1	SNN	B	147	7/8	0.98	0.07	14,14,16,16	0
1	SNN	C	147	7/8	0.98	0.07	14,14,15,16	0

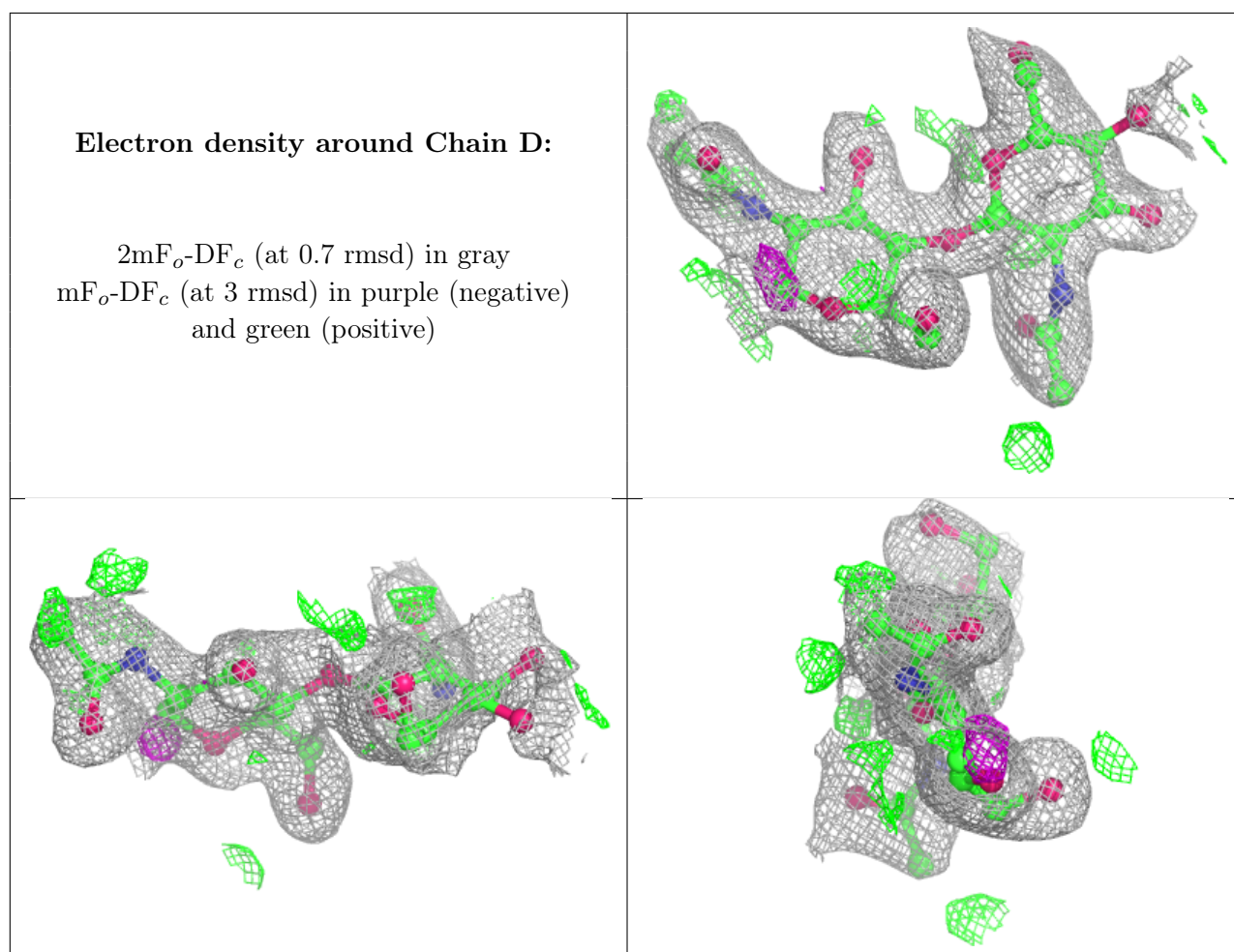
### 6.3 Carbohydrates [i](#)

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.

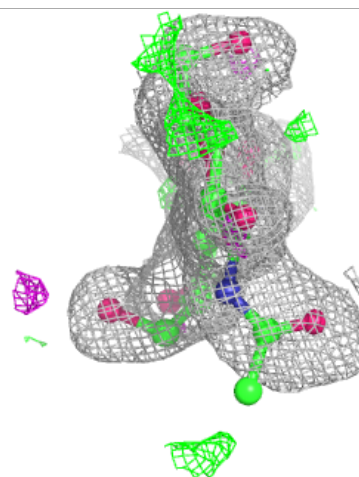
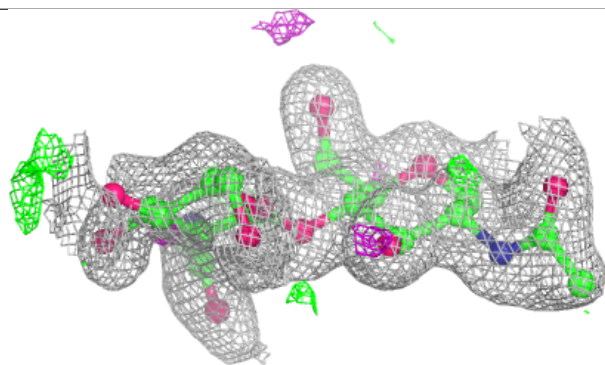
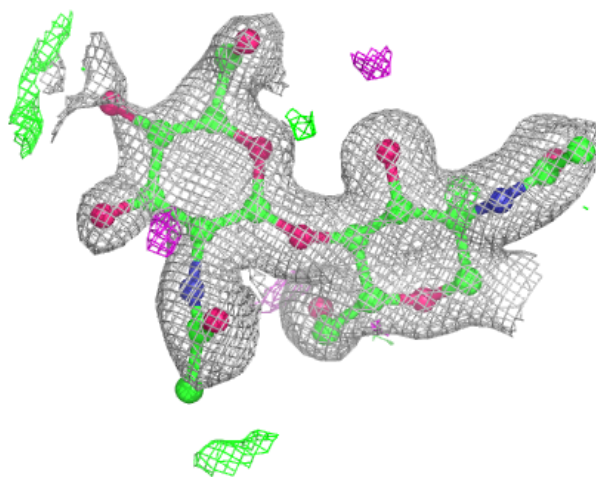
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAG	F	2	14/15	0.68	0.24	62,69,74,74	0
2	NAG	D	2	14/15	0.78	0.18	54,67,73,77	0
2	NAG	E	2	14/15	0.81	0.24	44,51,58,65	0
2	NAG	F	1	14/15	0.82	0.11	37,42,44,52	0
2	NAG	D	1	14/15	0.85	0.10	34,39,41,49	0
2	NAG	E	1	14/15	0.95	0.08	25,30,36,37	0

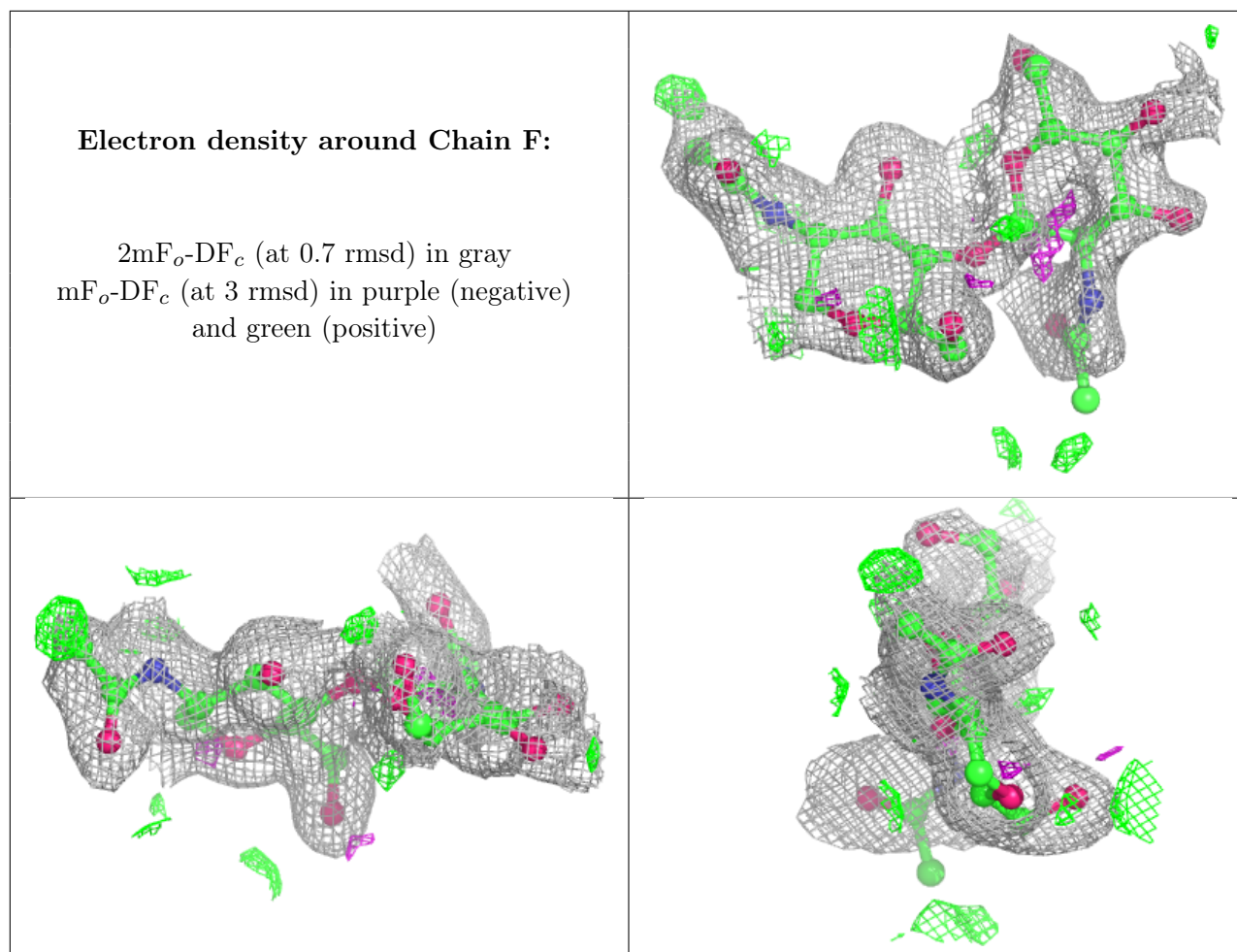
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**Electron density around Chain E:**

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





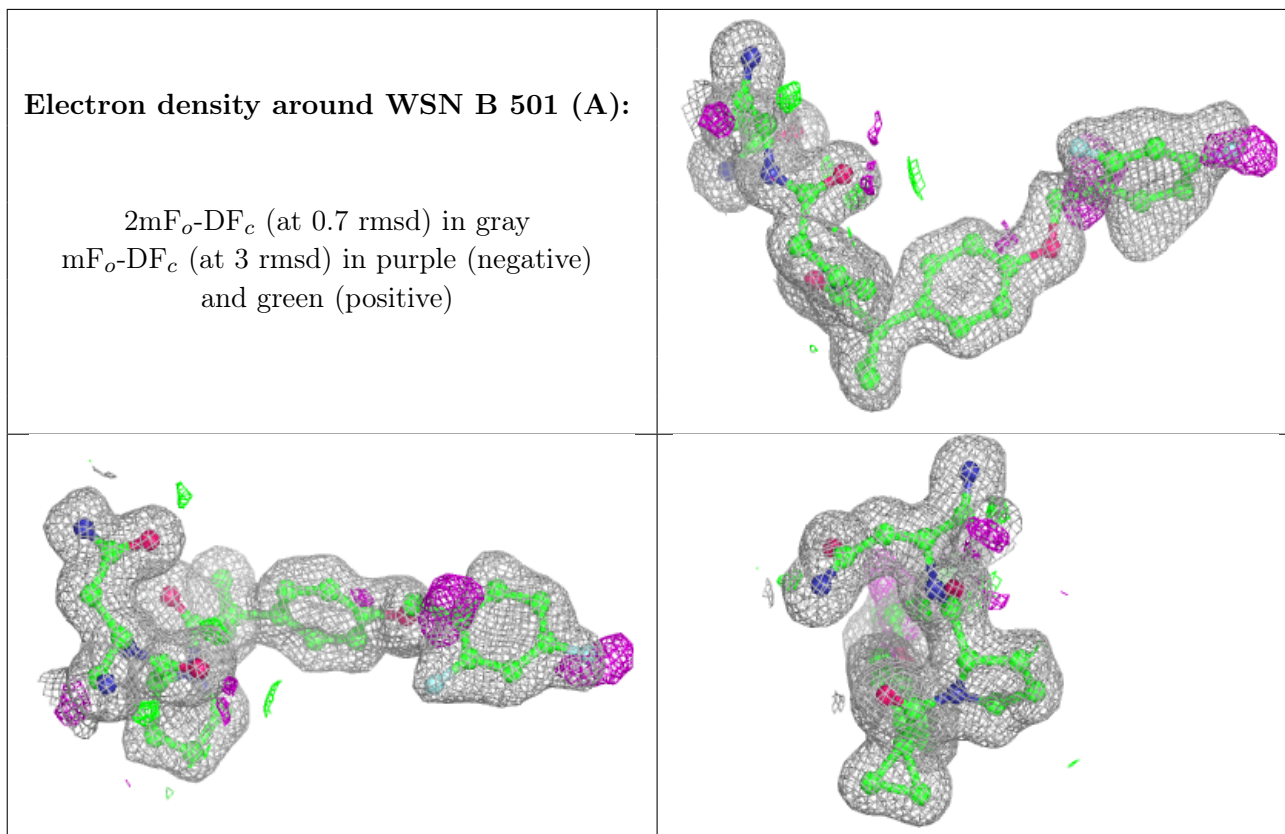
## 6.4 Ligands [i](#)

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
4	EDO	A	502	4/4	0.80	0.15	29,30,31,31	0
4	EDO	C	503	4/4	0.86	0.11	32,35,35,35	0
4	EDO	C	502	4/4	0.92	0.13	40,44,45,47	0
4	EDO	A	504	4/4	0.93	0.11	25,31,34,35	0
3	WSN	B	501[A]	36/36	0.95	0.10	15,20,33,38	9
3	WSN	B	501[B]	36/36	0.95	0.10	15,20,28,29	9
3	WSN	A	501	36/36	0.96	0.08	13,15,22,26	0
4	EDO	A	503	4/4	0.97	0.07	19,23,25,31	0
3	WSN	C	501	36/36	0.97	0.07	14,17,27,30	0

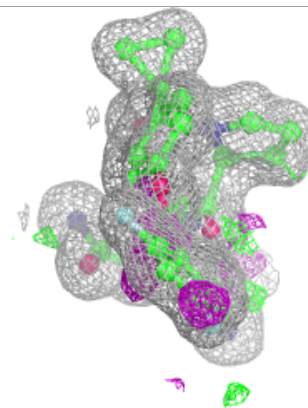
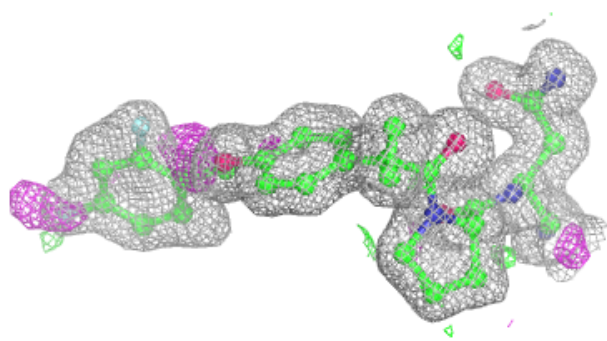
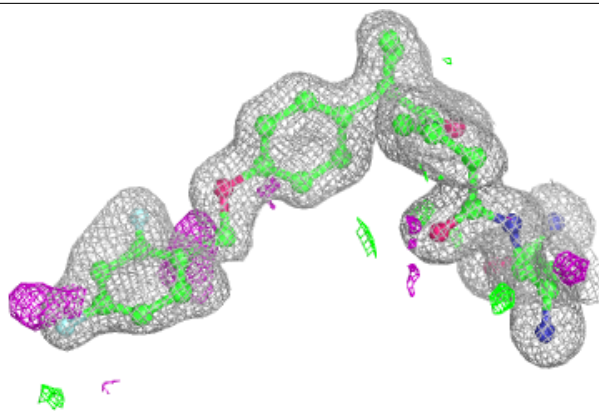
The following is a graphical depiction of the model fit to experimental electron density of all

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.

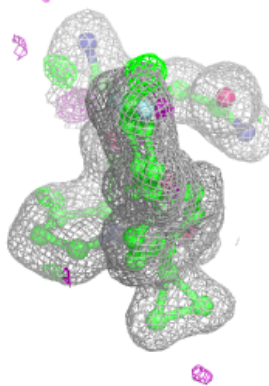
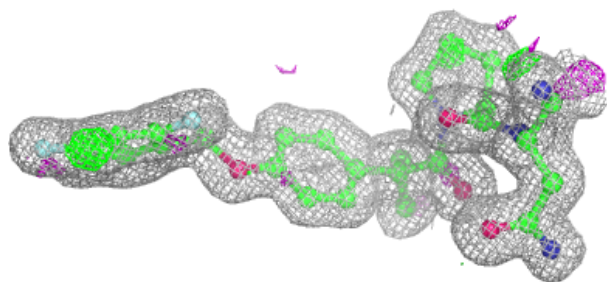
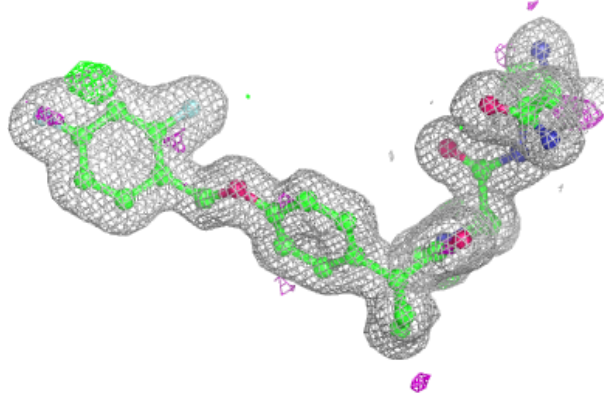


**Electron density around WSN B 501 (B):**

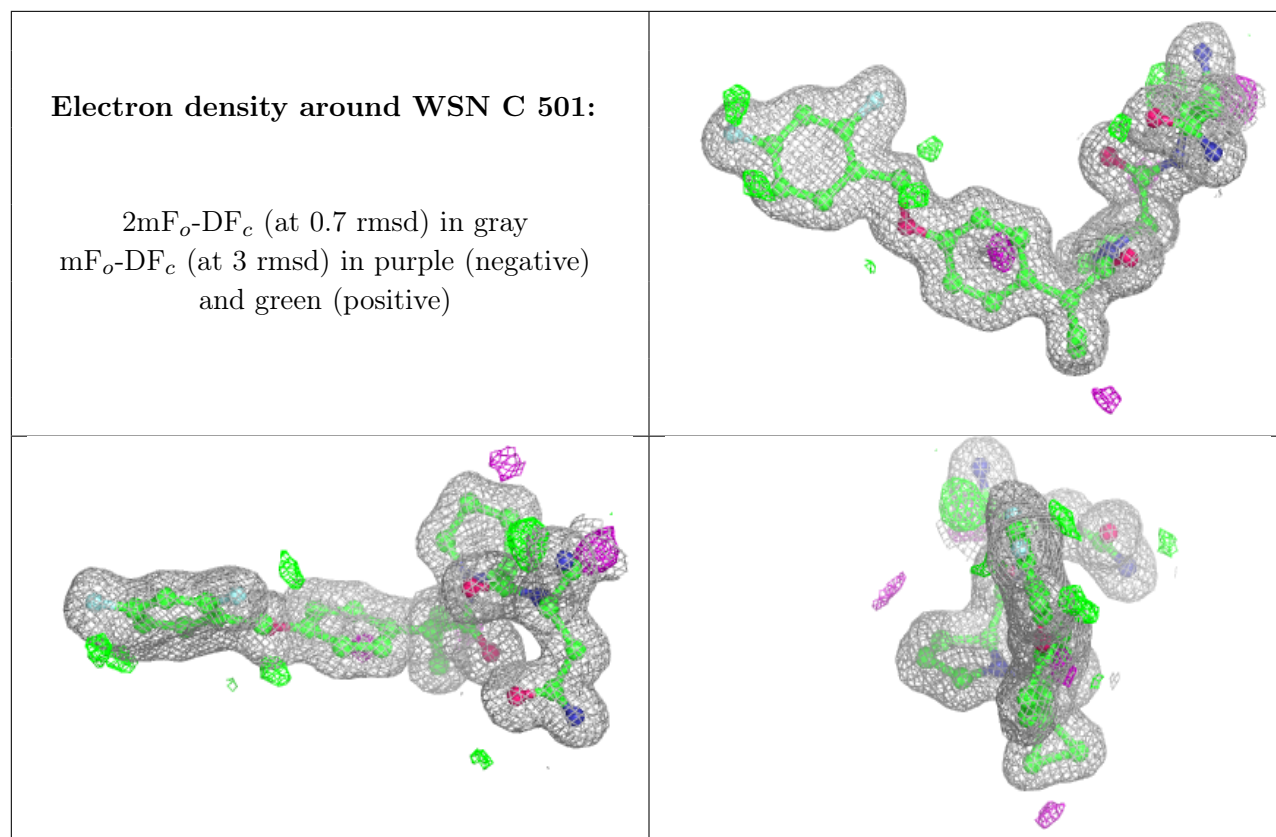
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around WSN A 501:**

$2mF_o-DF_c$  (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.