



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 14, 2024 – 08:33 AM EST

PDB ID : 7FQJ  
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]cyclopropa-1,2-dicarbonyl]pyrrolidine-2-carboxamide  
Authors : Ehler, A.; Benz, J.; Bartels, B.; Rudolph, M.G.  
Deposited on : 2022-10-05  
Resolution : 1.70 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	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

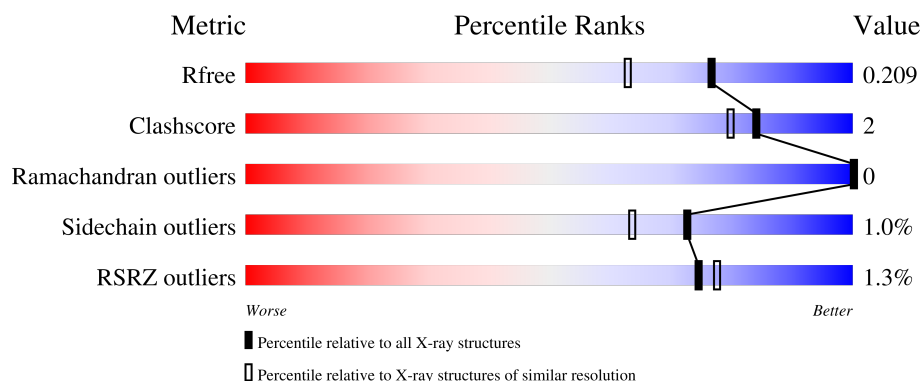
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	164625	5161 (1.70-1.70)
Clashscore	180529	5671 (1.70-1.70)
Ramachandran outliers	177936	5594 (1.70-1.70)
Sidechain outliers	177891	5594 (1.70-1.70)
RSRZ outliers	164620	5159 (1.70-1.70)

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	<div> <div>%</div> <div> <div></div> <div>55%</div> <div>5%</div> <div>40%</div> </div> </div>
1	B	444	<div> <div></div> <div>56%</div> <div>•</div> <div>41%</div> </div>
1	C	444	<div> <div>%</div> <div> <div></div> <div>55%</div> <div>•</div> <div>41%</div> </div> </div>
2	E	2	<div> <div></div> <div>50%</div> <div>50%</div> </div>

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 7348 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
1	A	265	Total	C	N	O	S	0	6	0
			2167	1371	372	409	15			
1	B	264	Total	C	N	O	S	0	6	0
			2162	1371	366	410	15			
1	C	260	Total	C	N	O	S	0	4	0
			2116	1341	360	400	15			

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.



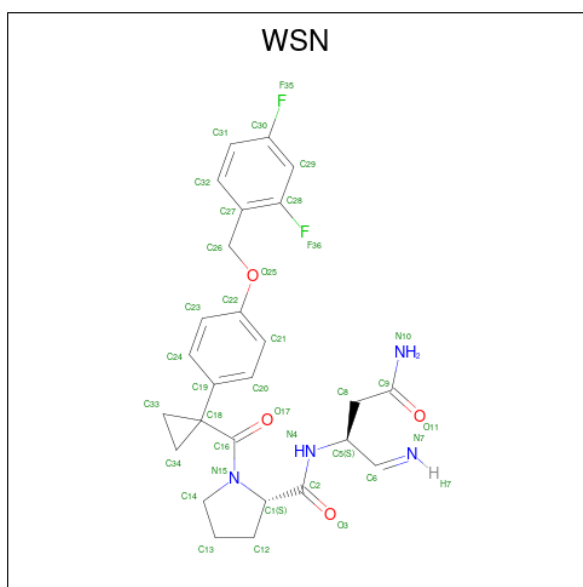
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	2	Total	C	N	O	0	0	0
			28	16	2	10			

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

- Molecule 4 is N-[(2R)-4-amino-1-imino-4-oxobutan-2-yl]-1-(1-{4-[(2,4-difluorophenyl)methoxy]phenyl}cyclopropane-1-carbonyl)-L-prolinamide (three-letter code: WSN) (formula: C<sub>26</sub>H<sub>28</sub>F<sub>2</sub>N<sub>4</sub>O<sub>4</sub>).



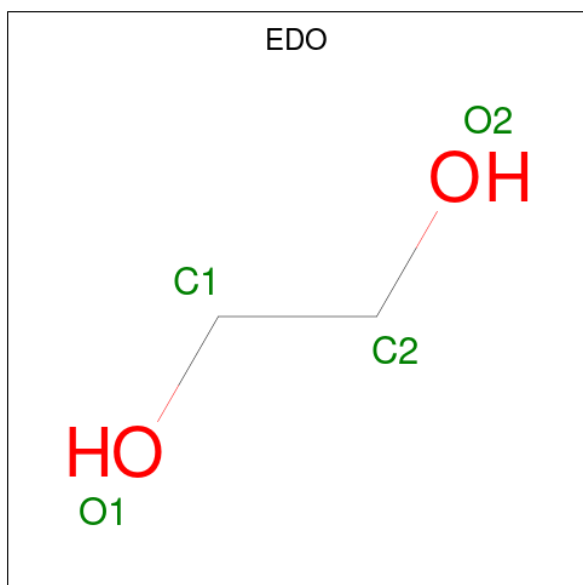
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	F	N	O	0
			36	26	2	4	4	

*Continued on next page...*

Continued from previous page...

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

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

- Molecule 6 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			13	8	5		

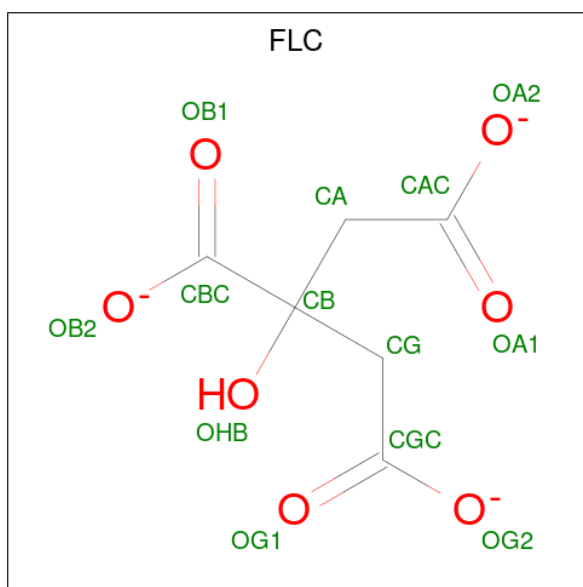
- Molecule 7 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula:  $C_8H_{18}N_2O_4S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 8 is CITRATE ANION (three-letter code: FLC) (formula:  $\text{C}_6\text{H}_5\text{O}_7$ ).



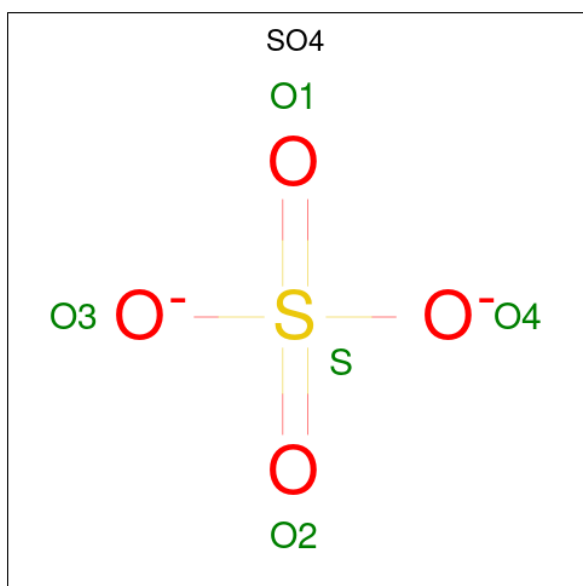


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	1	Total	C	O	0	0
			13	6	7		

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	B	1	Total	Mg	0	0
			1	1		

- Molecule 10 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	C	1	Total	O	S	0	0
			5	4	1		

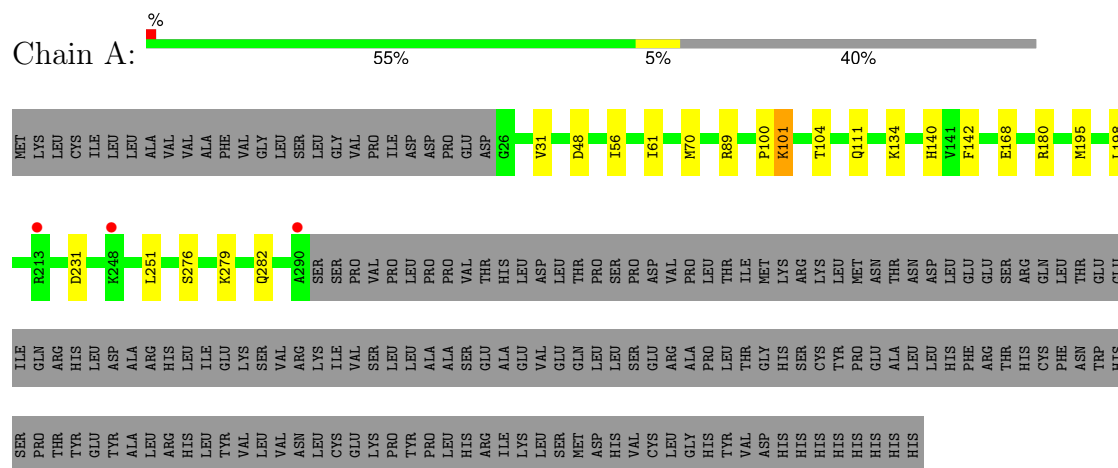
- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	245	Total	O	0	2
			247	247		
11	B	240	Total	O	0	1
			241	241		
11	C	179	Total	O	0	0
			179	179		

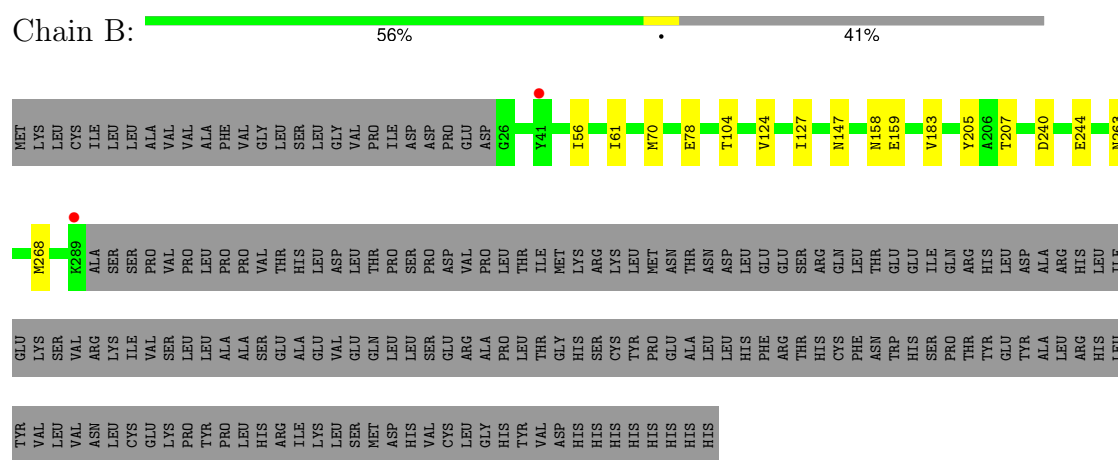
### 3 Residue-property plots [i](#)

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.

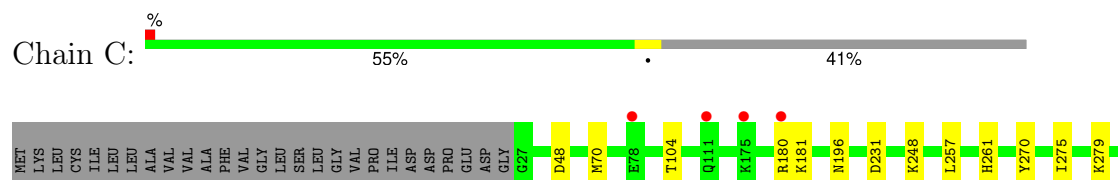
#### • Molecule 1: Legumain

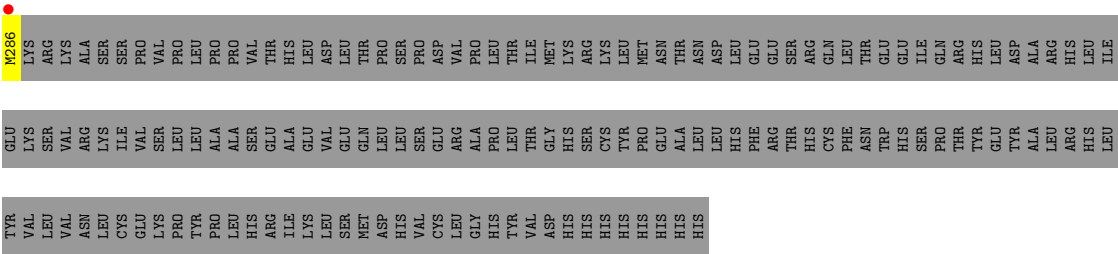


#### • Molecule 1: Legumain



#### • Molecule 1: Legumain





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



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	229.39Å 66.57Å 79.11Å 90.00° 100.93° 90.00°	Depositor
Resolution (Å)	112.62 – 1.70 112.62 – 1.70	Depositor EDS
% Data completeness (in resolution range)	70.2 (112.62-1.70) 70.6 (112.62-1.70)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.169 , 0.201 0.179 , 0.209	Depositor DCC
$R_{free}$ test set	6491 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.2	Xtriage
Anisotropy	0.038	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 31.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7348	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.74% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: FLC, EPE, WSN, SNN, SO4, MG, EDO, PG4, 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	0/2234	0.77	0/3026
1	B	0.64	0/2230	0.77	0/3019
1	C	0.68	0/2177	0.76	0/2953
All	All	0.66	0/6641	0.76	0/8998

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

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	2167	0	2074	15	0
1	B	2162	0	2065	6	0
1	C	2116	0	2016	8	0
2	E	28	0	25	0	0
3	A	14	0	13	0	0
3	C	14	0	13	0	0
4	A	36	0	0	0	0
4	B	45	0	0	0	0
4	C	36	0	0	0	0
5	A	4	0	6	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	12	0	18	2	0
6	A	13	0	18	0	0
7	A	15	0	17	0	0
8	B	13	0	4	0	0
9	B	1	0	0	0	0
10	C	5	0	0	0	0
11	A	247	0	0	6	0
11	B	241	0	0	0	0
11	C	179	0	0	0	0
All	All	7348	0	6269	29	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 29 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	11:A:601:HOH:O	1.77	1.00
1:A:89[A]:ARG:HD2	11:A:817[A]:HOH:O	1.92	0.69
1:A:111[B]:GLN:HG3	11:A:749:HOH:O	1.99	0.62
1:A:168[A]:GLU:HG2	11:A:635:HOH:O	2.05	0.56
1:B:240:ASP:OD1	1:B:244[B]:GLU:HG3	2.09	0.52

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	268/444 (60%)	262 (98%)	6 (2%)	0	100	100
1	B	267/444 (60%)	257 (96%)	10 (4%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	261/444 (59%)	254 (97%)	7 (3%)	0	100	100
All	All	796/1332 (60%)	773 (97%)	23 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	236/396 (60%)	234 (99%)	2 (1%)	79	71
1	B	236/396 (60%)	232 (98%)	4 (2%)	56	41
1	C	231/396 (58%)	230 (100%)	1 (0%)	89	85
All	All	703/1188 (59%)	696 (99%)	7 (1%)	73	64

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	158	ASN
1	B	159	GLU
1	C	286	MET
1	B	263	ASN
1	B	78	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	88	ASN
1	B	65	GLN
1	B	263	ASN
1	C	88	ASN
1	C	138	GLN



### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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	B	147	1	5,6,8	0.82	0	1,6,11	2.32	1 (100%)
1	SNN	C	147	1	5,6,8	0.82	0	1,6,11	1.01	0
1	SNN	A	147	1	5,6,8	0.84	0	1,6,11	1.73	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
1	SNN	B	147	1	-	3/3/5/12	-
1	SNN	C	147	1	-	3/3/5/12	-
1	SNN	A	147	1	-	3/3/5/12	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	147	SNN	O5-C5-C4	-2.32	118.64	125.38

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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
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

2 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	2,1	14,14,15	0.59	0	17,19,21	0.97	0
2	NAG	E	2	2	14,14,15	0.53	0	17,19,21	1.63	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	2,1	-	0/6/23/26	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	2	NAG	C1-O5-C5	5.80	119.97	112.19

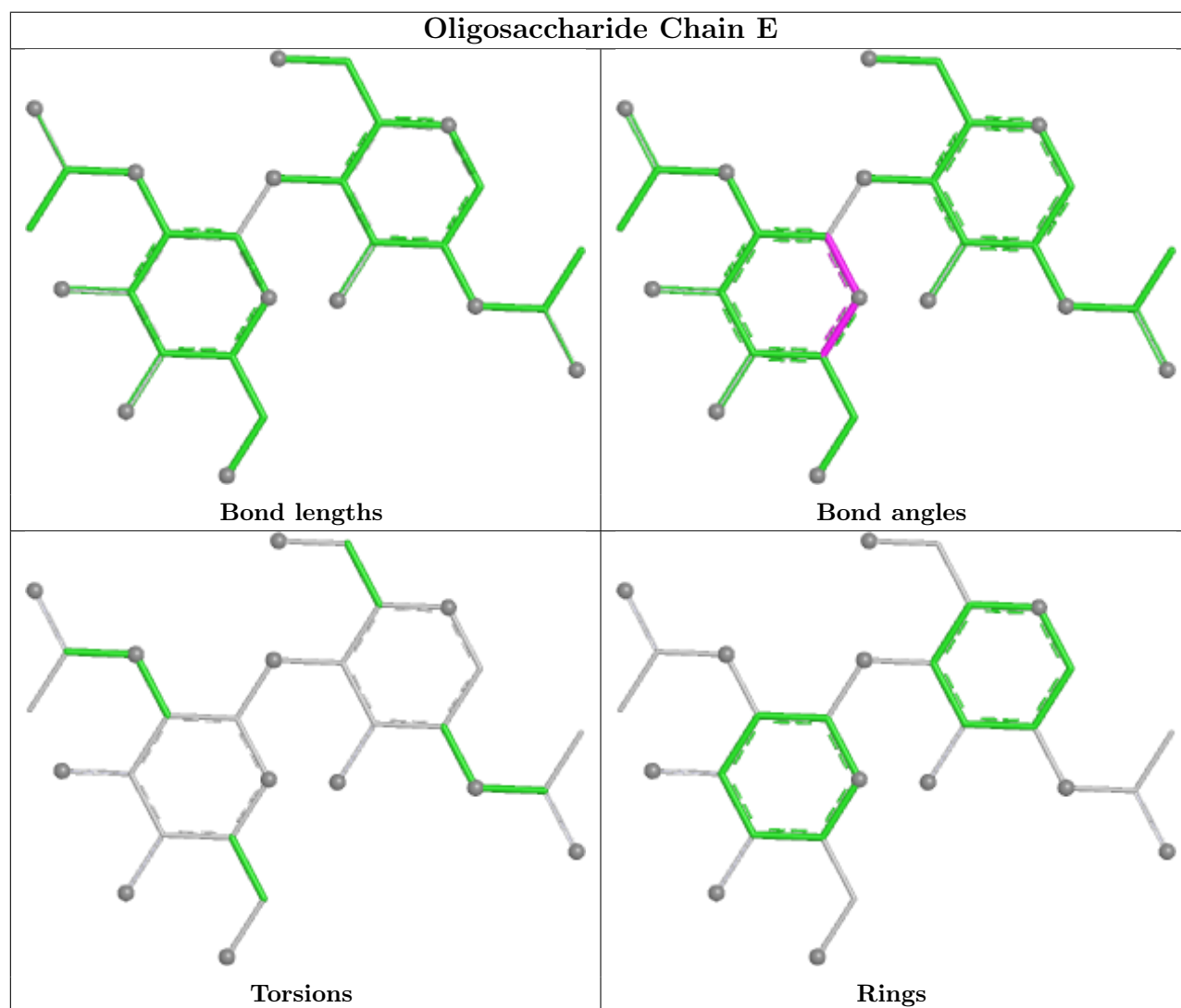
There are no chirality outliers.

There are no torsion outliers.

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](#)

Of 15 ligands modelled in this entry, 1 is monoatomic - leaving 14 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	$\# Z  > 2$
4	WSN	B	501[B]	-	37,39,39	0.35	0	52,56,56	0.83	3 (5%)
3	NAG	A	501	1	14,14,15	0.59	0	17,19,21	0.76	0
4	WSN	B	501[A]	-	37,39,39	0.33	0	52,56,56	0.77	2 (3%)
4	WSN	C	502	-	37,39,39	0.31	0	52,56,56	0.87	1 (1%)
10	SO4	C	506	-	4,4,4	0.27	0	6,6,6	0.17	0
3	NAG	C	501	1	14,14,15	0.35	0	17,19,21	0.81	1 (5%)
6	PG4	A	504	-	12,12,12	0.18	0	11,11,11	0.08	0
7	EPE	A	505	-	15,15,15	0.95	1 (6%)	19,20,20	1.45	2 (10%)
4	WSN	A	502	-	37,39,39	0.37	0	52,56,56	0.83	2 (3%)
5	EDO	A	503	-	3,3,3	0.20	0	2,2,2	0.19	0
5	EDO	C	504	-	3,3,3	0.11	0	2,2,2	0.21	0
5	EDO	C	505	-	3,3,3	0.08	0	2,2,2	0.29	0
8	FLC	B	502	9	12,12,12	1.30	1 (8%)	17,17,17	1.30	2 (11%)
5	EDO	C	503	-	3,3,3	0.18	0	2,2,2	0.31	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	WSN	B	501[B]	-	-	6/33/49/49	0/4/4/4
3	NAG	A	501	1	-	0/6/23/26	0/1/1/1
4	WSN	B	501[A]	-	-	6/33/49/49	0/4/4/4
4	WSN	C	502	-	-	4/33/49/49	0/4/4/4
3	NAG	C	501	1	-	0/6/23/26	0/1/1/1
6	PG4	A	504	-	-	3/10/10/10	-
7	EPE	A	505	-	-	3/9/19/19	0/1/1/1
4	WSN	A	502	-	-	0/33/49/49	0/4/4/4
5	EDO	A	503	-	-	0/1/1/1	-
5	EDO	C	504	-	-	0/1/1/1	-
5	EDO	C	505	-	-	1/1/1/1	-
8	FLC	B	502	9	-	0/16/16/16	-
5	EDO	C	503	-	-	0/1/1/1	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	505	EPE	O1S-S	3.55	1.55	1.45
8	B	502	FLC	CB-CBC	2.04	1.55	1.53

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	505	EPE	O1S-S-C10	-4.68	99.65	106.73
7	A	505	EPE	O3S-S-O2S	3.86	121.05	111.40
4	C	502	WSN	C33-C18-C19	3.58	122.76	118.58
4	B	501[A]	WSN	C33-C18-C19	3.43	122.58	118.58
4	B	501[B]	WSN	C33-C18-C19	3.43	122.58	118.58

There are no chirality outliers.

5 of 23 torsion outliers are listed below:

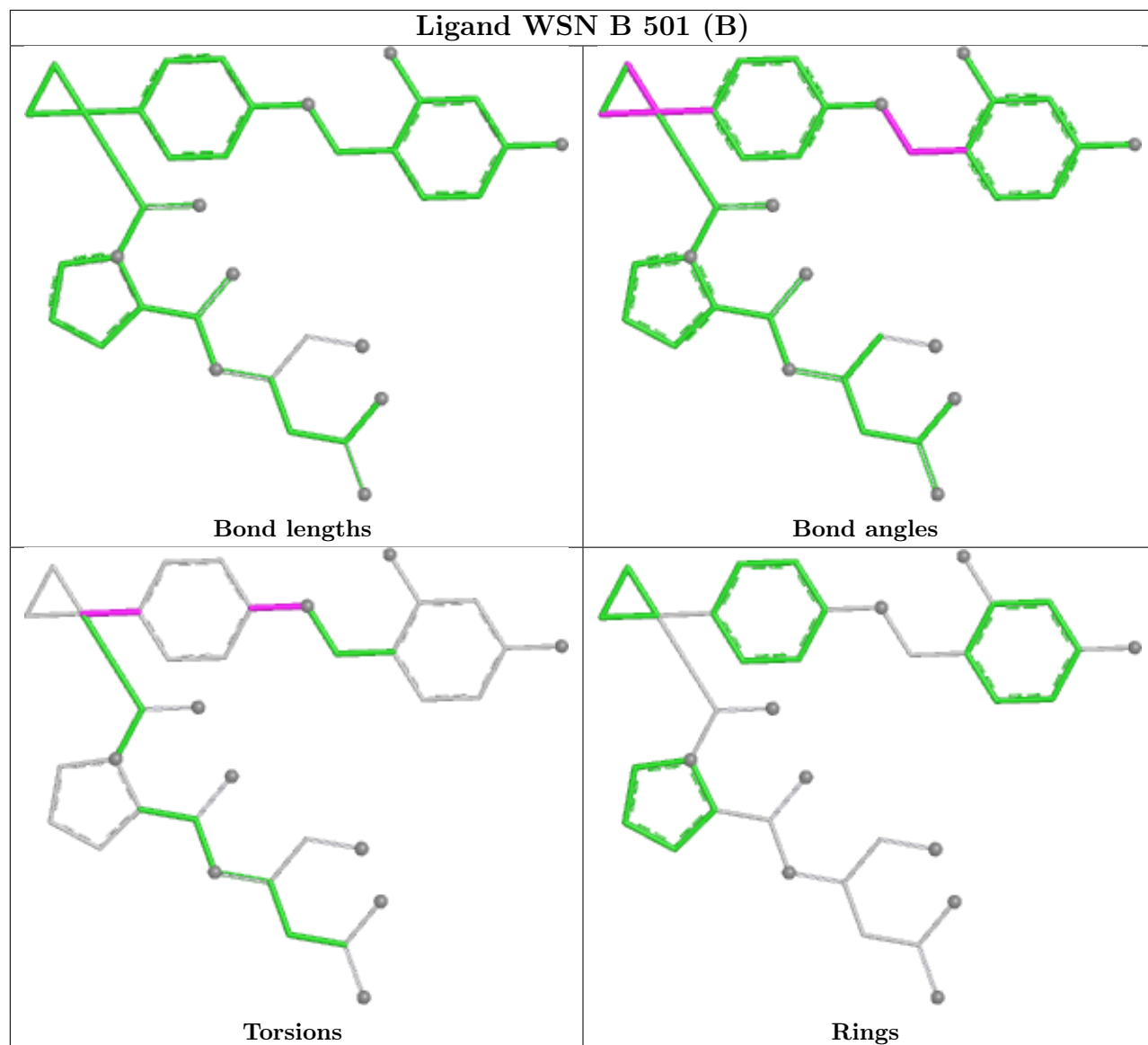
Mol	Chain	Res	Type	Atoms
4	C	502	WSN	C16-C18-C19-C24
7	A	505	EPE	C9-C10-S-O2S
7	A	505	EPE	C9-C10-S-O3S
4	B	501[A]	WSN	C23-C22-O25-C26
4	B	501[A]	WSN	C21-C22-O25-C26

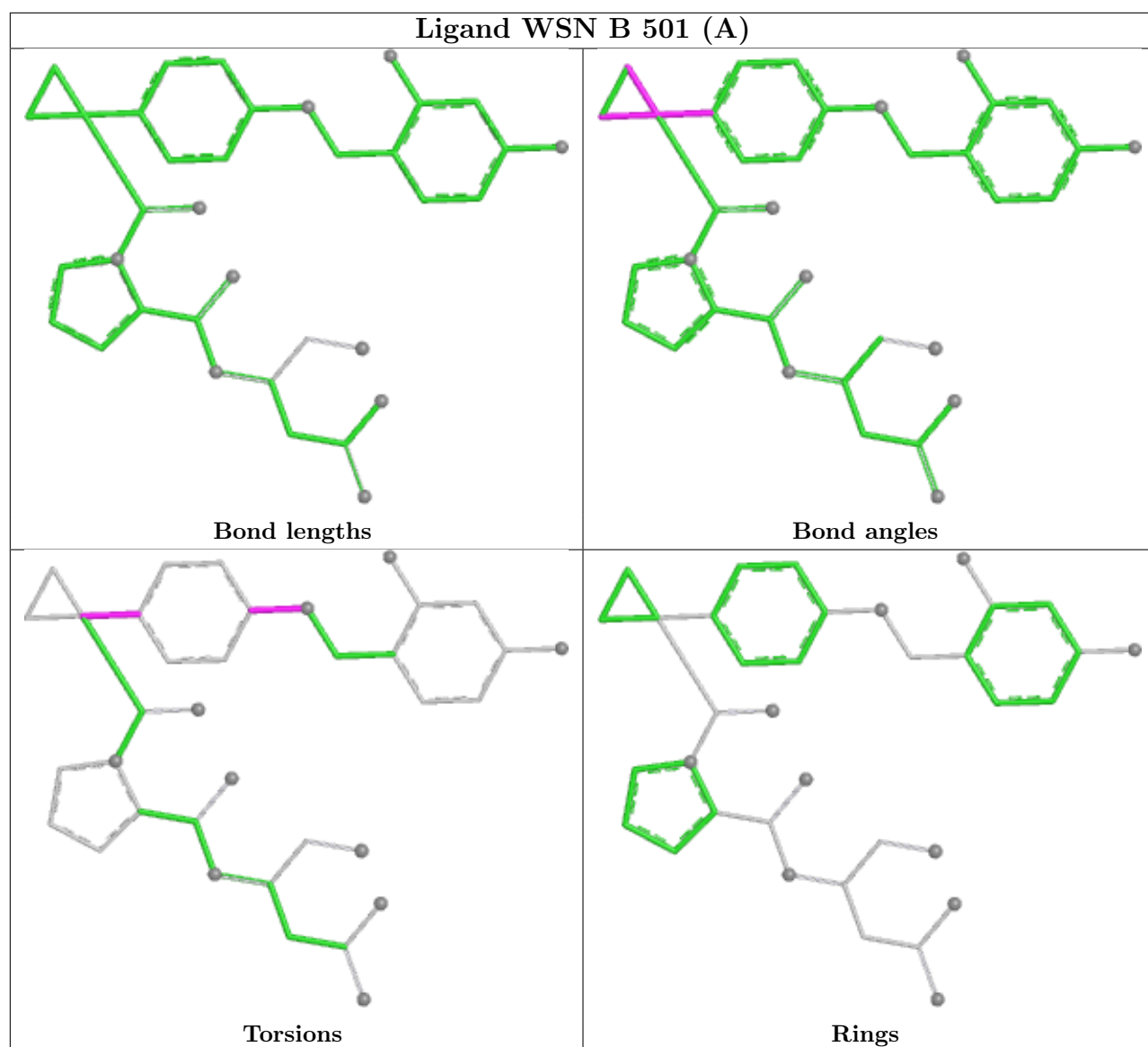
There are no ring outliers.

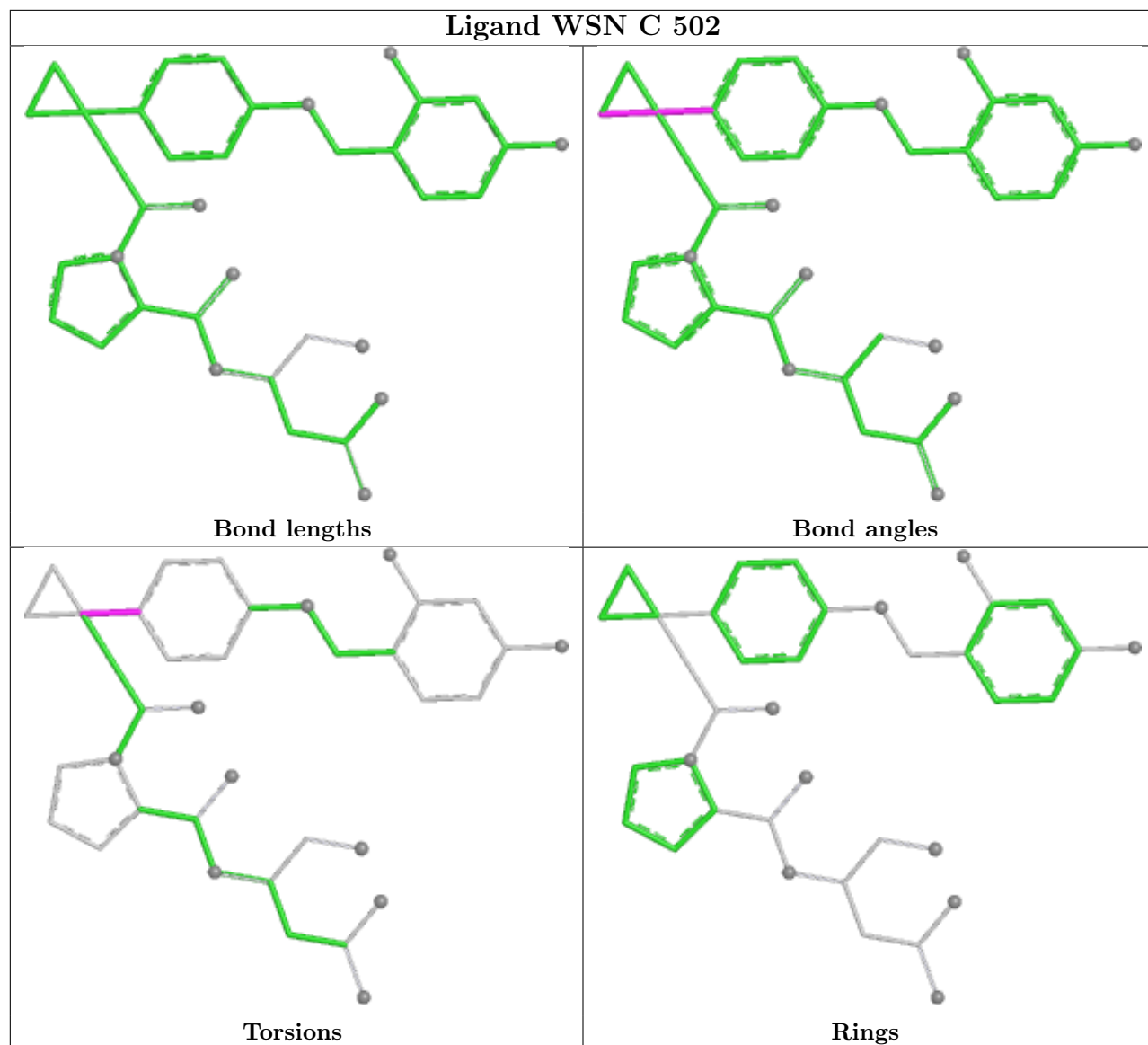
1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	504	EDO	2	0

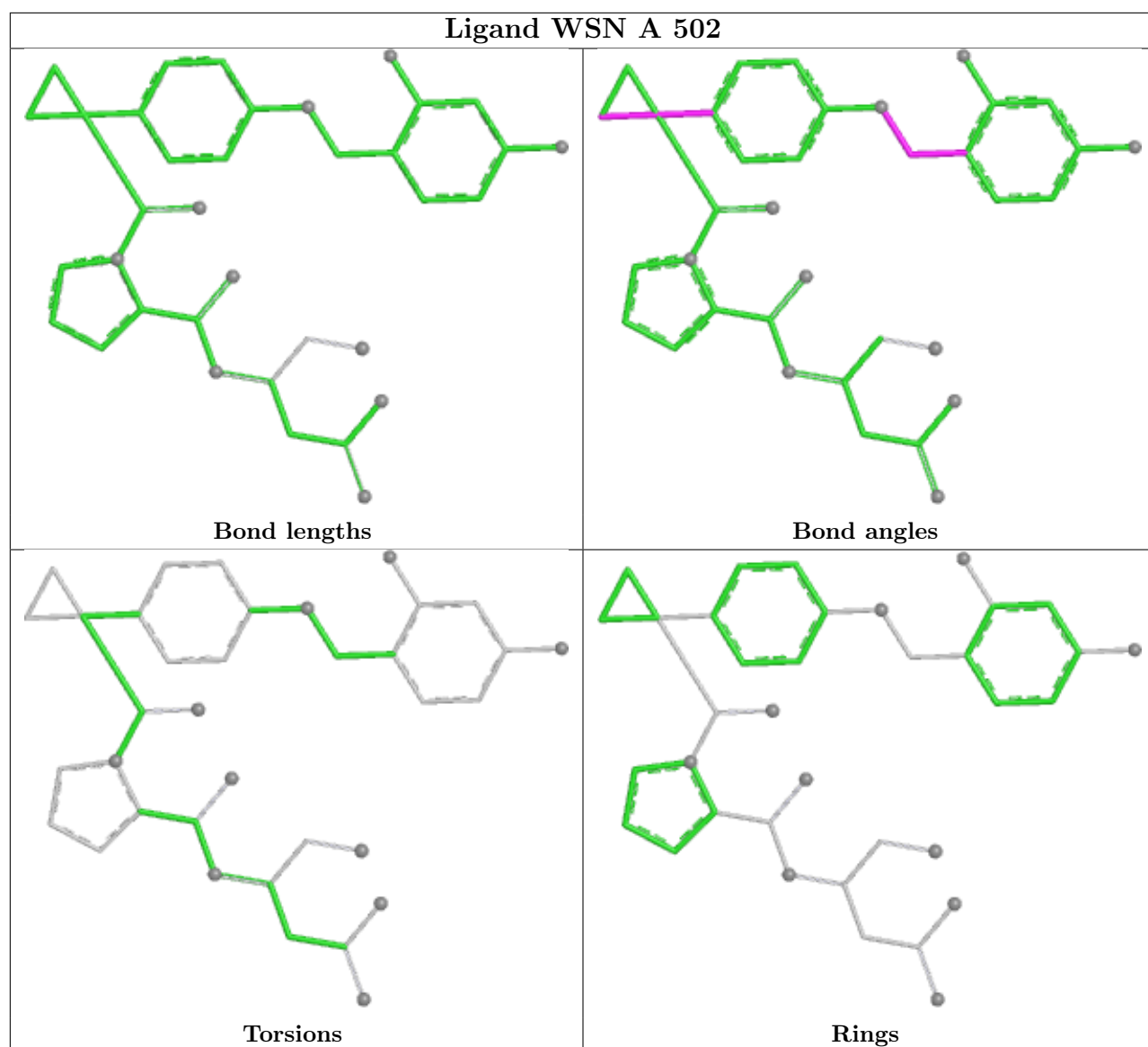
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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











## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	264/444 (59%)	-0.25	3 (1%) 77 80	8, 17, 33, 49	6 (2%)
1	B	263/444 (59%)	-0.25	2 (0%) 82 85	12, 18, 31, 62	6 (2%)
1	C	259/444 (58%)	-0.02	5 (1%) 66 69	9, 21, 40, 57	4 (1%)
All	All	786/1332 (59%)	-0.17	10 (1%) 74 77	8, 19, 34, 62	16 (2%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	286	MET	5.2
1	B	289	LYS	4.8
1	A	290	ALA	4.5
1	C	175	LYS	3.3
1	C	78	GLU	2.8

### 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	0.05	11,11,12,13	0
1	SNN	B	147	7/8	0.98	0.04	12,13,15,15	0
1	SNN	C	147	7/8	0.98	0.05	13,14,15,16	0

### 6.3 Carbohydrates [i](#)

SUGAR-RSR INFOmissingINFO

## 6.4 Ligands

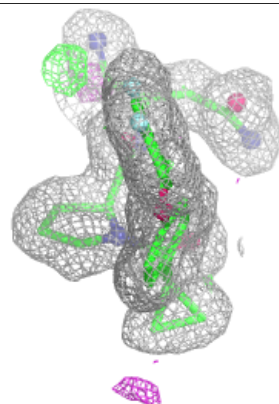
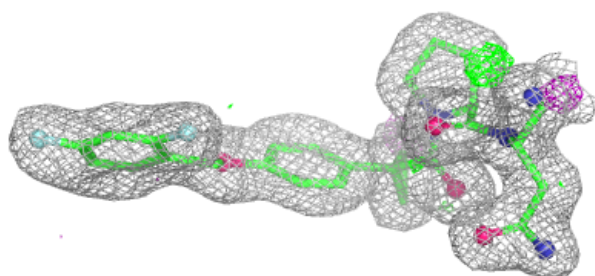
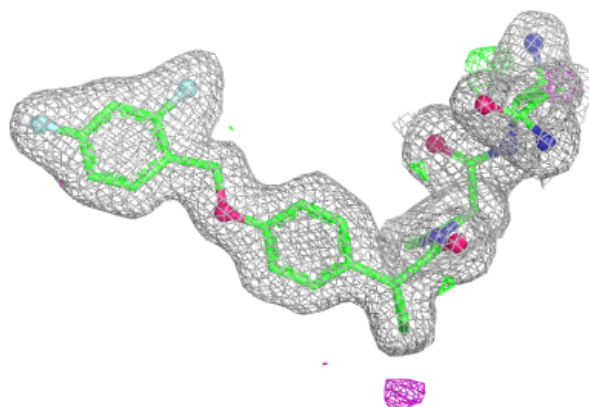
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<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
5	EDO	C	505	4/4	0.76	0.20	44,45,47,49	0
7	EPE	A	505	15/15	0.80	0.22	48,55,63,63	0
3	NAG	A	501	14/15	0.81	0.13	38,41,42,49	0
10	SO4	C	506	5/5	0.83	0.15	54,64,66,70	0
8	FLC	B	502	13/13	0.85	0.12	29,33,37,38	0
5	EDO	A	503	4/4	0.85	0.16	38,41,41,41	0
5	EDO	C	503	4/4	0.88	0.14	32,35,38,41	0
5	EDO	C	504	4/4	0.89	0.14	38,40,41,49	0
3	NAG	C	501	14/15	0.90	0.10	35,37,41,48	0
6	PG4	A	504	13/13	0.92	0.15	43,44,48,49	13
9	MG	B	503	1/1	0.94	0.27	32,32,32,32	0
4	WSN	C	502	36/36	0.96	0.06	14,17,25,27	0
4	WSN	A	502	36/36	0.96	0.06	12,15,25,31	0
4	WSN	B	501[A]	36/36	0.97	0.07	16,19,39,43	9
4	WSN	B	501[B]	36/36	0.97	0.07	16,19,27,29	9

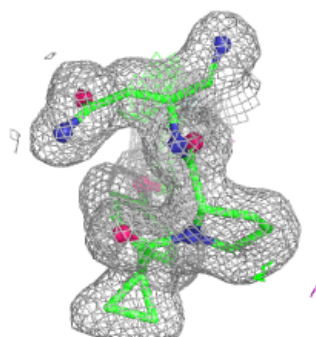
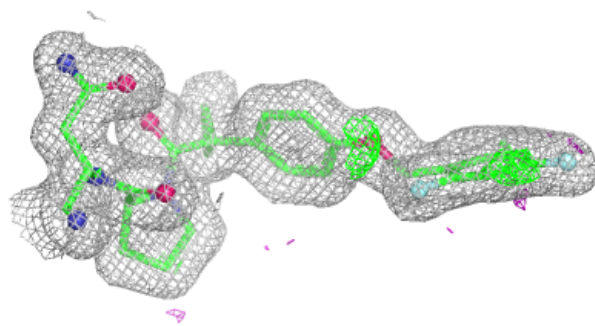
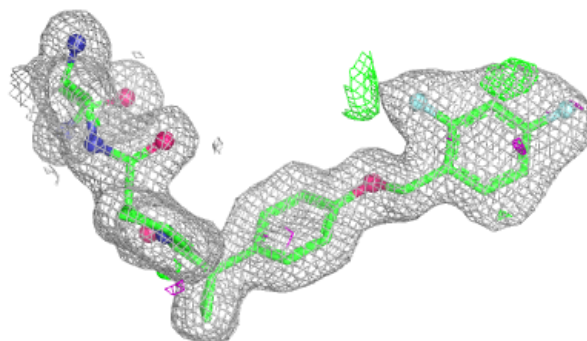
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around WSN C 502:**

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

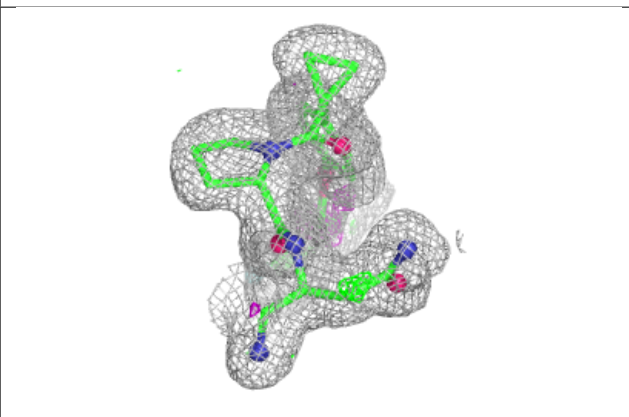
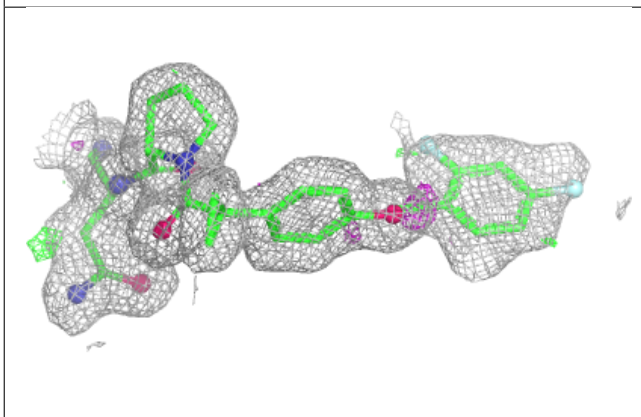
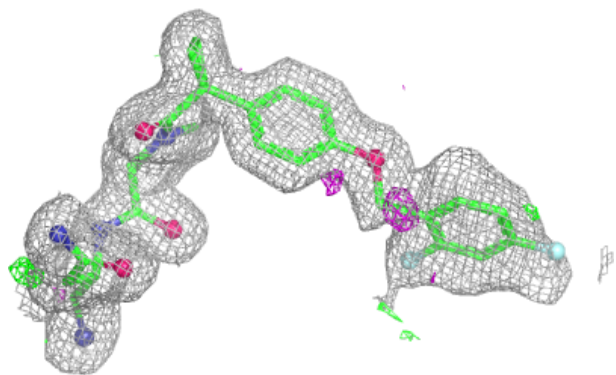
**Electron density around WSN A 502:**

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

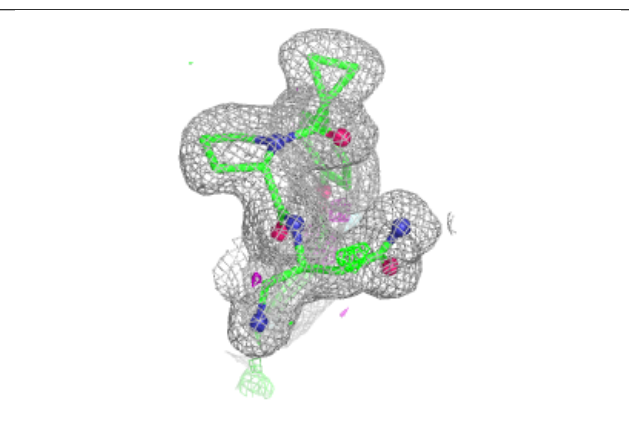
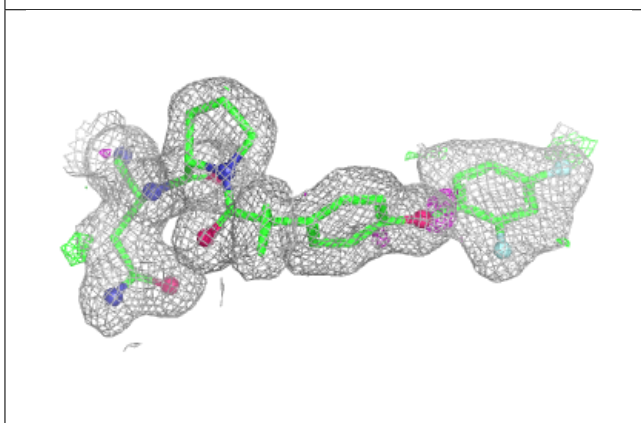
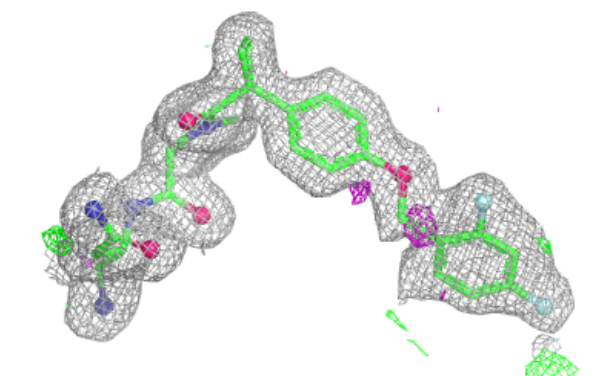


**Electron density around WSN B 501 (A):**

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

**Electron density around WSN 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)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.