



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 22, 2023 – 06:23 PM EDT

PDB ID : 2ZZL  
Title : Structure of bacteriorhodopsin's M intermediate at pH 7  
Authors : Yamamoto, M.; Kouyama, T.  
Deposited on : 2009-02-18  
Resolution : 2.03 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

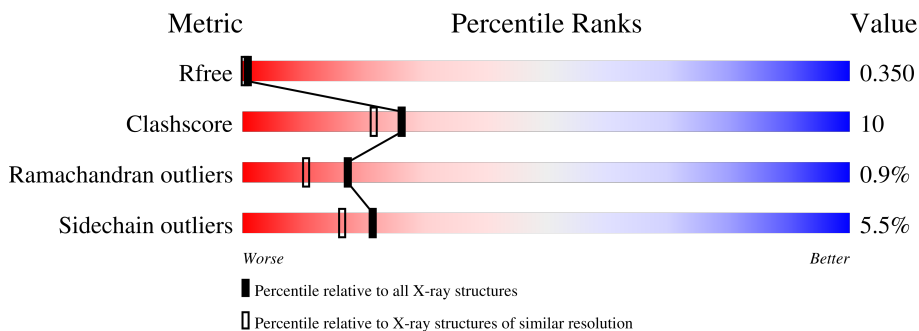
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.03 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	10434 (2.04-2.00)
Clashscore	141614	11643 (2.04-2.00)
Ramachandran outliers	138981	11493 (2.04-2.00)
Sidechain outliers	138945	11492 (2.04-2.00)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	262	 68% 16% 13%
2	B	3	 33% 67%

## 2 Entry composition [i](#)

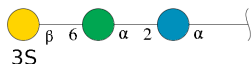
There are 8 unique types of molecules in this entry. The entry contains 2127 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Bacteriorhodopsin.

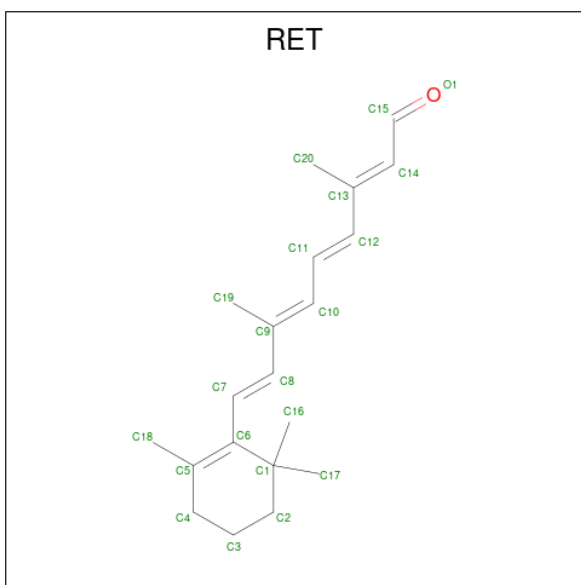
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	228	1765	1185	269	302	9	0	0	0

- Molecule 2 is an oligosaccharide called 3-O-sulfo-beta-D-galactopyranose-(1-6)-alpha-D-mannopyranose-(1-2)-alpha-D-glucopyranose.



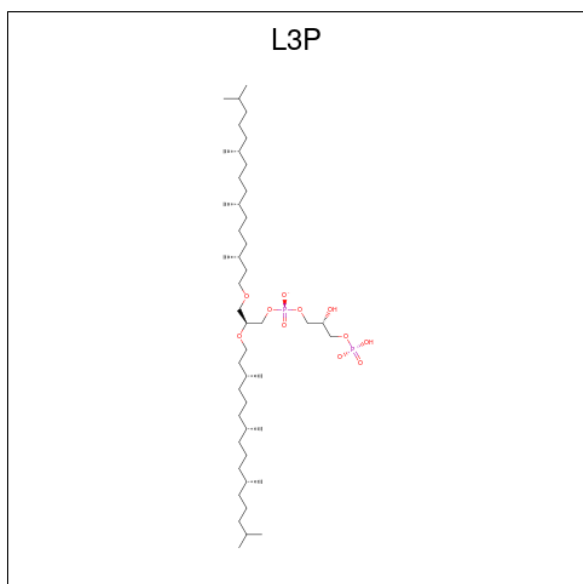
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	O	S			
2	B	3	37	18	18	1	0	0	0

- Molecule 3 is RETINAL (three-letter code: RET) (formula: C<sub>20</sub>H<sub>28</sub>O).



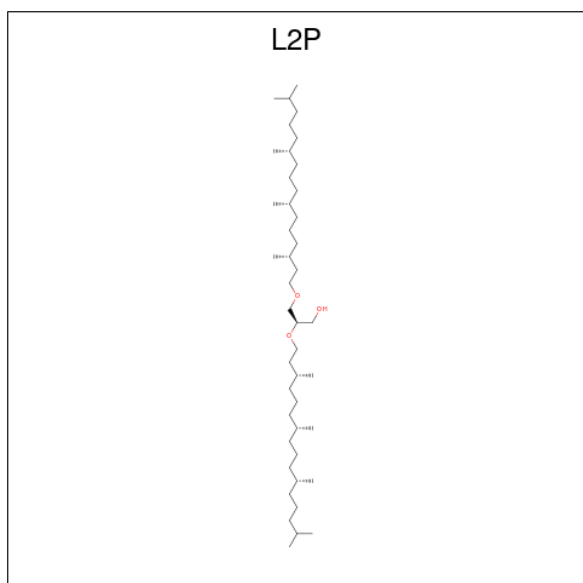
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C 20 20	0	0

- Molecule 4 is 2,3-DI-O-PHYTANLY-3-SN-GLYCERO-1-PHOSPHORYL-3'-SN-GLYCEROL-1'-PHOSPHATE (three-letter code: L3P) (formula:  $C_{46}H_{94}O_{11}P_2$ ).



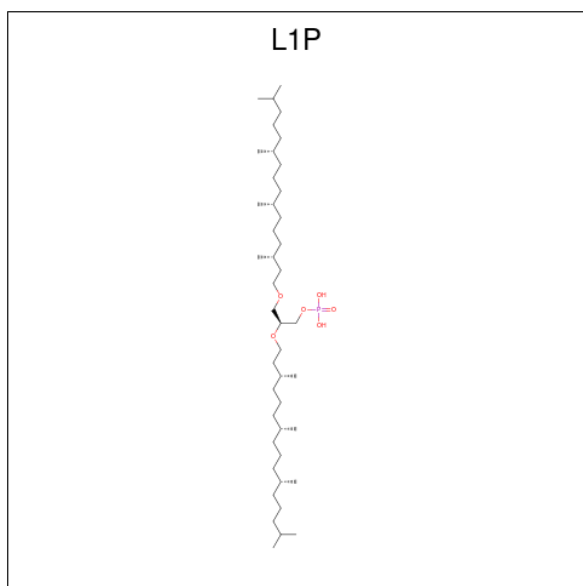
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O P 59 46 11 2	0	0

- Molecule 5 is 2,3-DI-PHYTANYL-GLYCEROL (three-letter code: L2P) (formula:  $C_{43}H_{88}O_3$ ).



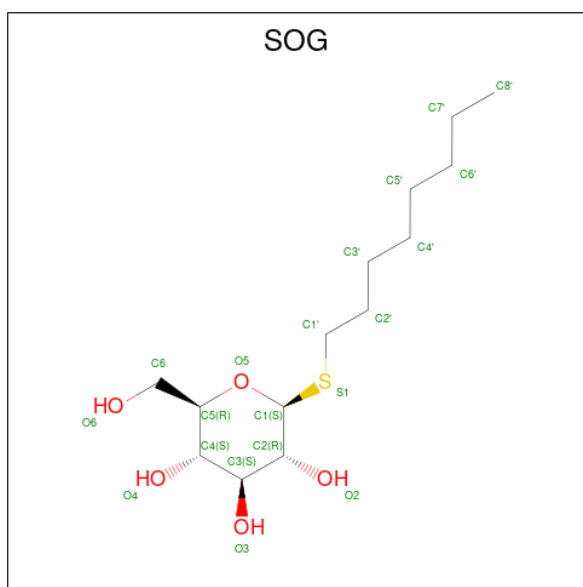
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	46	43	3	0	0

- Molecule 6 is 3-PHOSPHORYL-[1,2-DI-PHYTANYL]GLYCEROL (three-letter code: L1P) (formula:  $C_{43}H_{89}O_6P$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
6	A	1	40	33	6	1	0	0
6	A	1	50	43	6	1	0	0
6	A	1	50	43	6	1	0	0

- Molecule 7 is octyl 1-thio-beta-D-glucopyranoside (three-letter code: SOG) (formula:  $C_{14}H_{28}O_5S$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	O	S	0	0
			20	14	5	1		

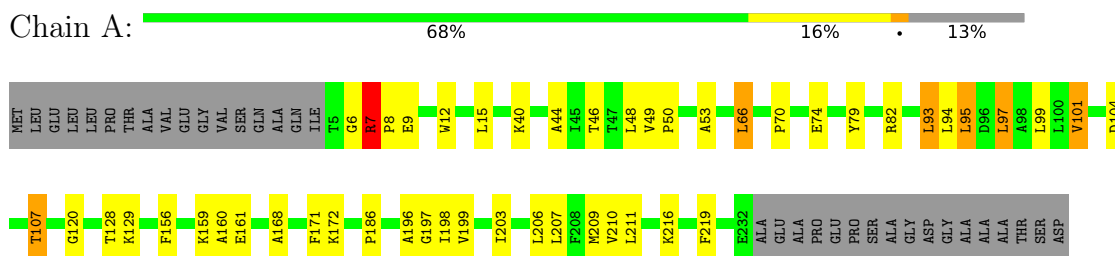
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	40	Total	O	0	0
			40	40		

### 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. 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. 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: Bacteriorhodopsin



- Molecule 2: 3-O-sulfo-beta-D-galactopyranose-(1-6)-alpha-D-mannopyranose-(1-2)-alpha-D-glucopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 6 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.21Å 102.21Å 112.48Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	15.00 – 2.03 41.18 – 2.03	Depositor EDS
% Data completeness (in resolution range)	88.3 (15.00-2.03) 93.8 (41.18-2.03)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.33 (at 2.03Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.207 , 0.238 0.361 , 0.350	Depositor DCC
$R_{free}$ test set	1054 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.9	Xtrriage
Anisotropy	0.259	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 73.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	2127	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.79% 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: L2P, L1P, MAN, SOG, GLC, L3P, SGA, RET

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	0.37	0/1813	0.53	0/2476

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	1765	0	1819	45	0
2	B	37	0	27	3	0
3	A	20	0	27	2	0
4	A	59	0	93	5	0
5	A	46	0	87	1	0
6	A	140	0	238	0	0
7	A	20	0	28	3	0
8	A	40	0	0	1	0
All	All	2127	0	2319	46	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 46 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:129:LYS:HG3	7:A:430:SOG:H62	1.67	0.77
1:A:101:VAL:HG22	1:A:159:LYS:HB3	1.69	0.74
1:A:7:ARG:H	1:A:8:PRO:HD2	1.53	0.71
1:A:101:VAL:CG2	1:A:159:LYS:HB3	2.21	0.71
1:A:6:GLY:HA2	1:A:9:GLU:OE1	1.95	0.67

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	226/262 (86%)	220 (97%)	4 (2%)	2 (1%)	17 10

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	7	ARG
1	A	197	GLY

### 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	183/206 (89%)	173 (94%)	10 (6%)	21 16

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	107	THR
1	A	161	GLU
1	A	211	LEU
1	A	94	LEU
1	A	95	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	105	GLN
1	A	202	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](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

3 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	GLC	B	1	2	11,11,12	2.97	5 (45%)	15,15,17	1.09	1 (6%)
2	MAN	B	2	2	11,11,12	2.20	6 (54%)	15,15,17	1.75	4 (26%)
2	SGA	B	3	2	15,15,16	1.21	2 (13%)	17,22,24	1.15	2 (11%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GLC	B	1	2	-	0/2/19/22	0/1/1/1
2	MAN	B	2	2	-	2/2/19/22	0/1/1/1
2	SGA	B	3	2	-	0/7/24/27	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	GLC	O5-C5	5.22	1.54	1.43
2	B	1	GLC	O5-C1	5.04	1.51	1.43
2	B	1	GLC	C1-C2	4.56	1.62	1.52
2	B	2	MAN	C2-C3	4.03	1.58	1.52
2	B	1	GLC	C2-C3	3.85	1.58	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2	MAN	C1-O5-C5	4.08	117.73	112.19
2	B	2	MAN	O5-C1-C2	-3.05	106.07	110.77
2	B	2	MAN	O6-C6-C5	2.89	121.20	111.29
2	B	3	SGA	C3-O3-S	-2.70	113.66	118.88
2	B	3	SGA	C2-C3-C4	-2.44	108.53	110.66

There are no chirality outliers.

All (2) torsion outliers are listed below:

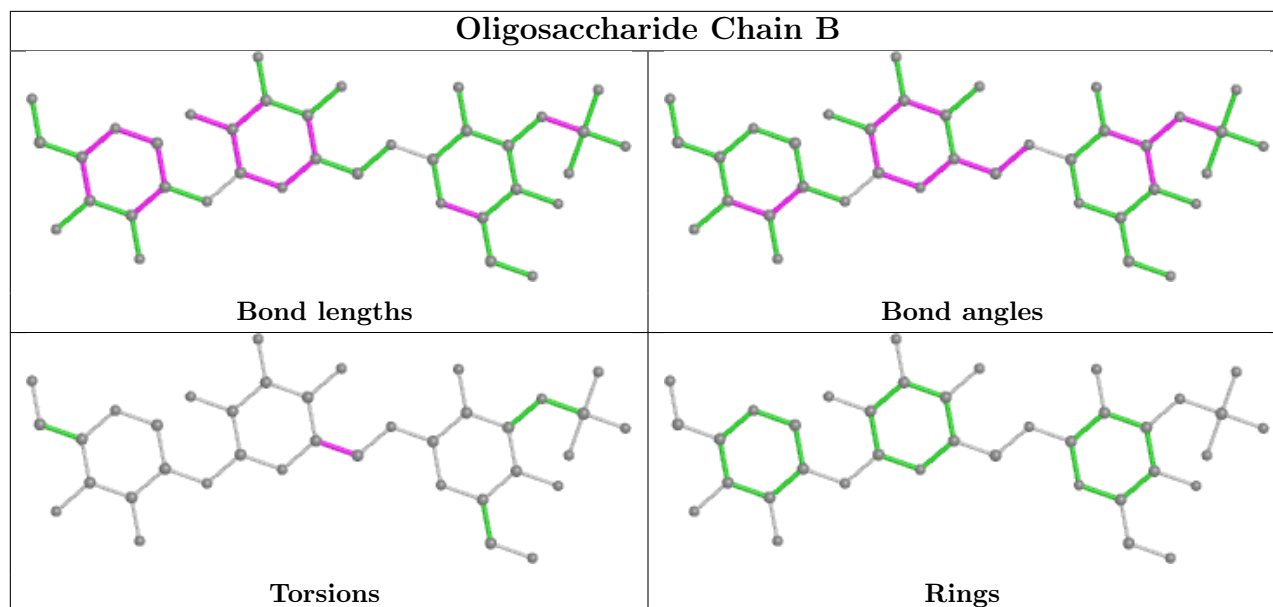
Mol	Chain	Res	Type	Atoms
2	B	2	MAN	O5-C5-C6-O6
2	B	2	MAN	C4-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	3	SGA	1	0
2	B	2	MAN	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 oligosaccharide.



## 5.6 Ligand geometry [i](#)

7 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$
7	SOG	A	430	-	20,20,20	1.47	4 (20%)	24,25,25	1.21	3 (12%)
4	L3P	A	260	-	58,58,58	1.22	8 (13%)	67,73,73	1.27	7 (10%)
6	L1P	A	290	-	49,49,49	1.78	15 (30%)	57,60,60	1.25	6 (10%)
5	L2P	A	270	-	45,45,45	1.87	13 (28%)	51,53,53	1.58	8 (15%)
6	L1P	A	280	-	39,39,49	1.90	12 (30%)	45,48,60	1.03	2 (4%)
6	L1P	A	300	-	49,49,49	1.66	12 (24%)	57,60,60	1.08	6 (10%)
3	RET	A	250	-	20,20,21	1.65	4 (20%)	27,27,28	2.11	11 (40%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	SOG	A	430	-	-	3/11/31/31	0/1/1/1
4	L3P	A	260	-	-	32/67/67/67	-
6	L1P	A	290	-	-	25/55/55/55	-
5	L2P	A	270	-	-	20/51/51/51	-
6	L1P	A	280	-	-	18/43/43/55	-
6	L1P	A	300	-	-	30/55/55/55	-
3	RET	A	250	-	-	0/13/30/31	0/1/1/1

The worst 5 of 68 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	280	L1P	O2-C2	4.97	1.57	1.43
6	A	290	L1P	P-O3	4.34	1.74	1.60
6	A	300	L1P	P-O3	4.18	1.73	1.60
5	A	270	L2P	C26-C25	4.16	1.70	1.52
6	A	280	L1P	P-O3P	4.02	1.70	1.54

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	290	L1P	O1-C1-C2	5.19	121.65	109.44
4	A	260	L3P	C41-O2-C2	4.99	126.86	115.40
5	A	270	L2P	C21-C22-C23	4.11	129.22	115.92
3	A	250	RET	C2-C1-C6	4.04	116.70	110.48
3	A	250	RET	C20-C13-C14	-3.94	112.05	123.71

There are no chirality outliers.

5 of 128 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	260	L3P	C4-O4-P1-O1P
4	A	260	L3P	C4-O4-P1-O2P
4	A	260	L3P	C41-C42-C43-C44
6	A	280	L1P	C49-C48-C50-C51
6	A	290	L1P	C3-O3-P-O1P

There are no ring outliers.

4 monomers are involved in 11 short contacts:

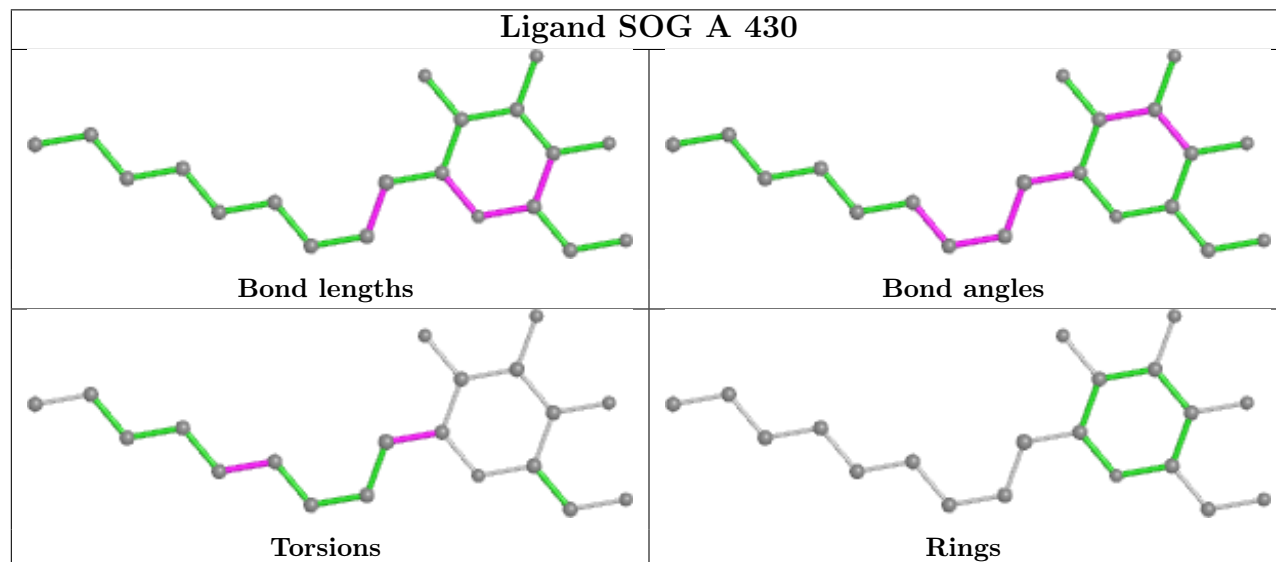
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	430	SOG	3	0

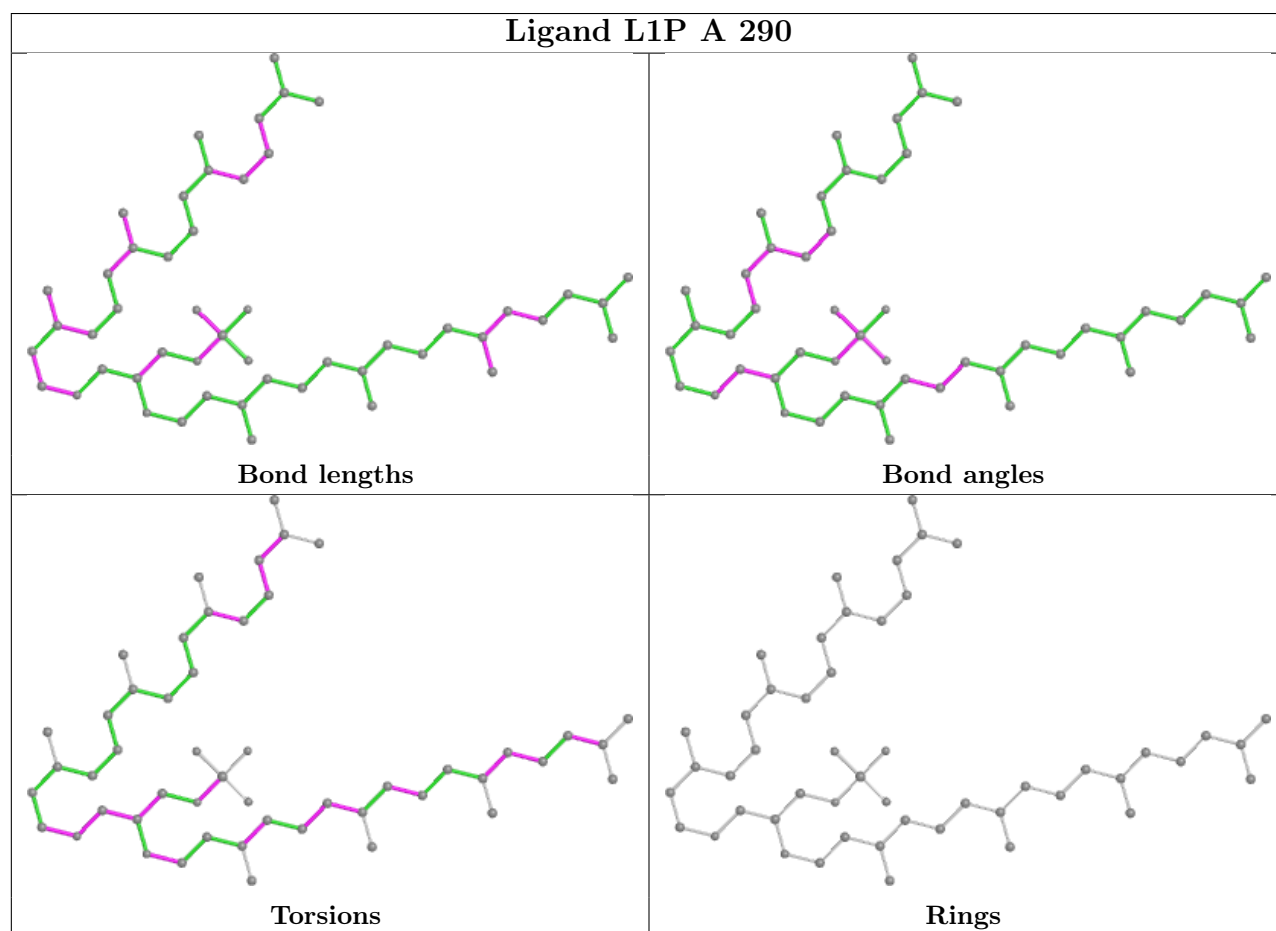
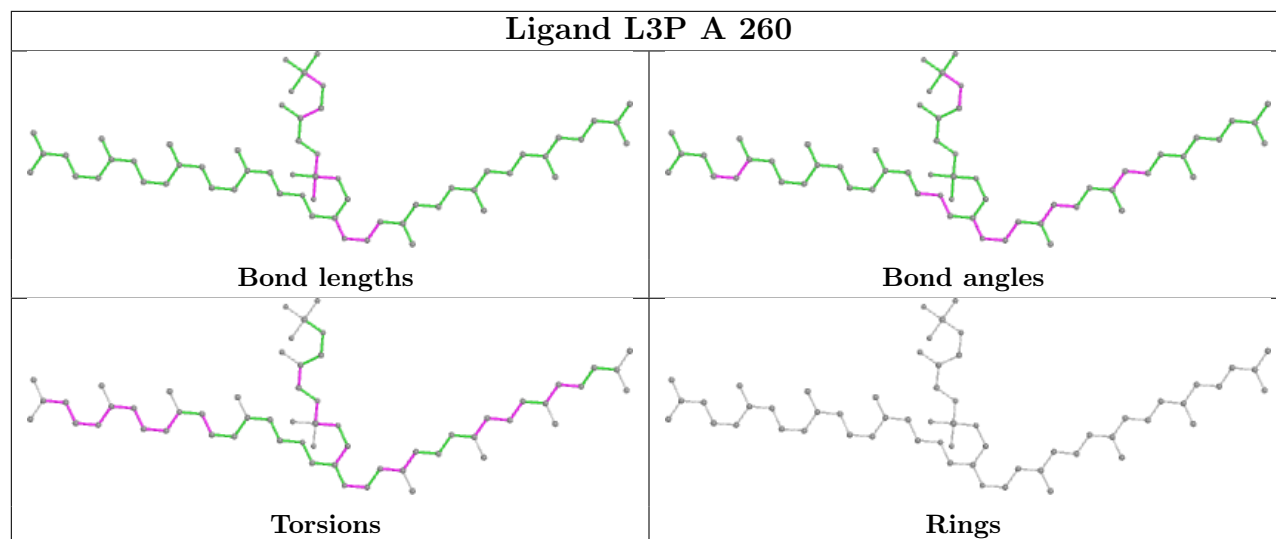
*Continued on next page...*

*Continued from previous page...*

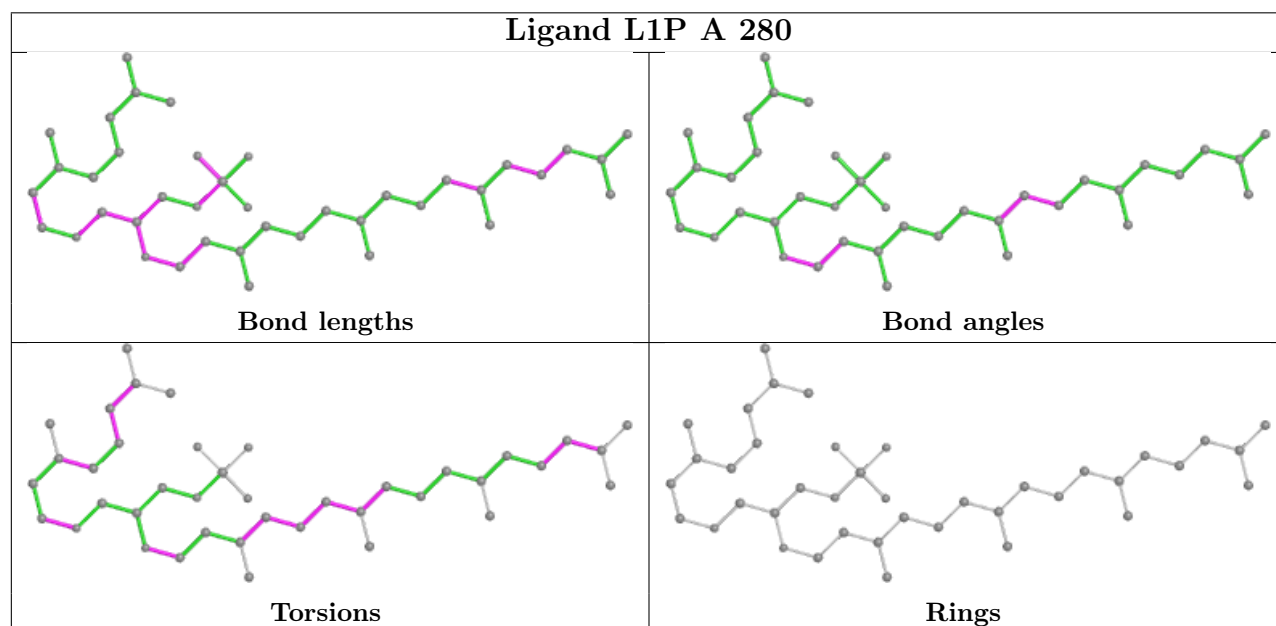
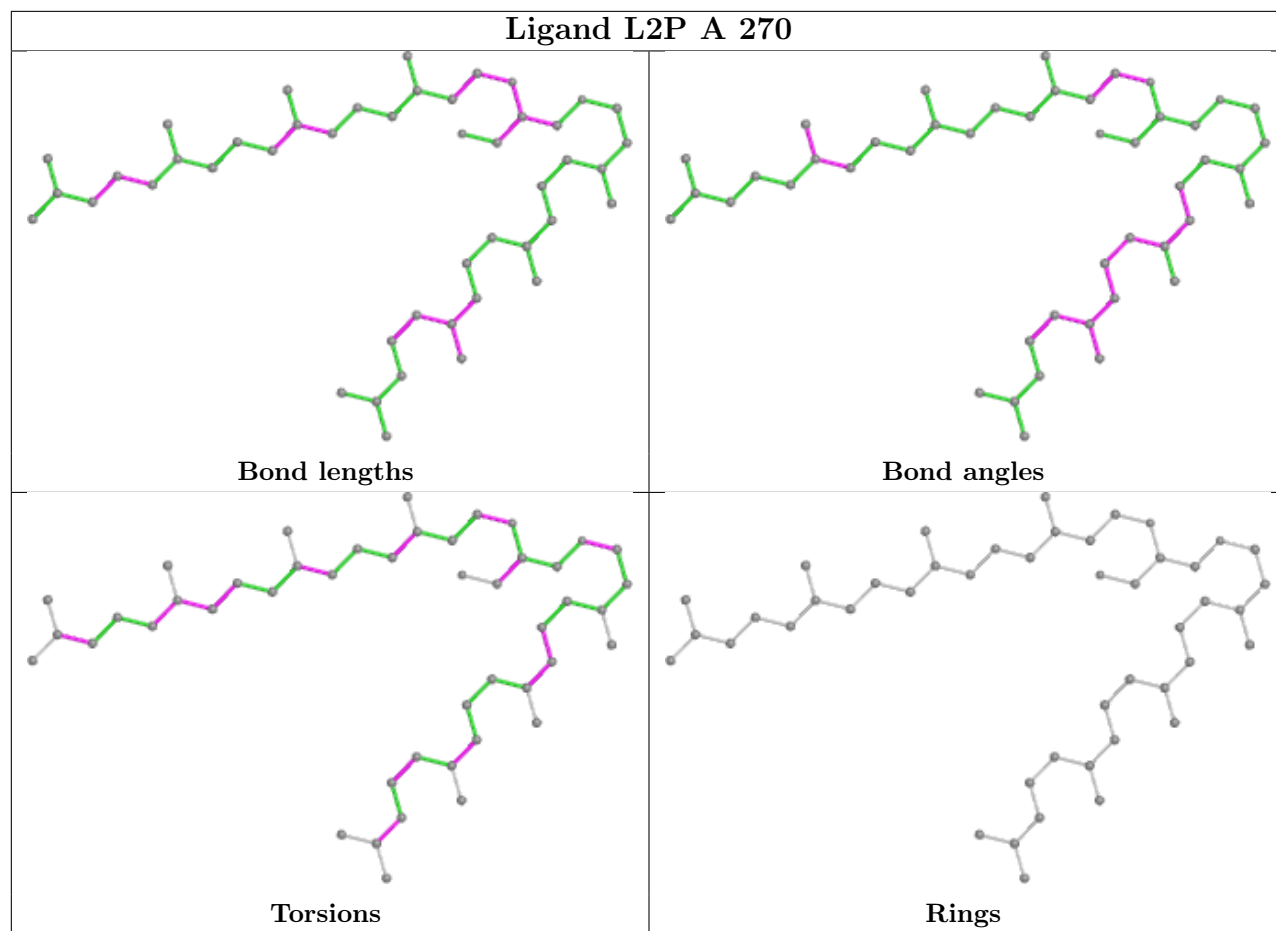
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	260	L3P	5	0
5	A	270	L2P	1	0
3	A	250	RET	2	0

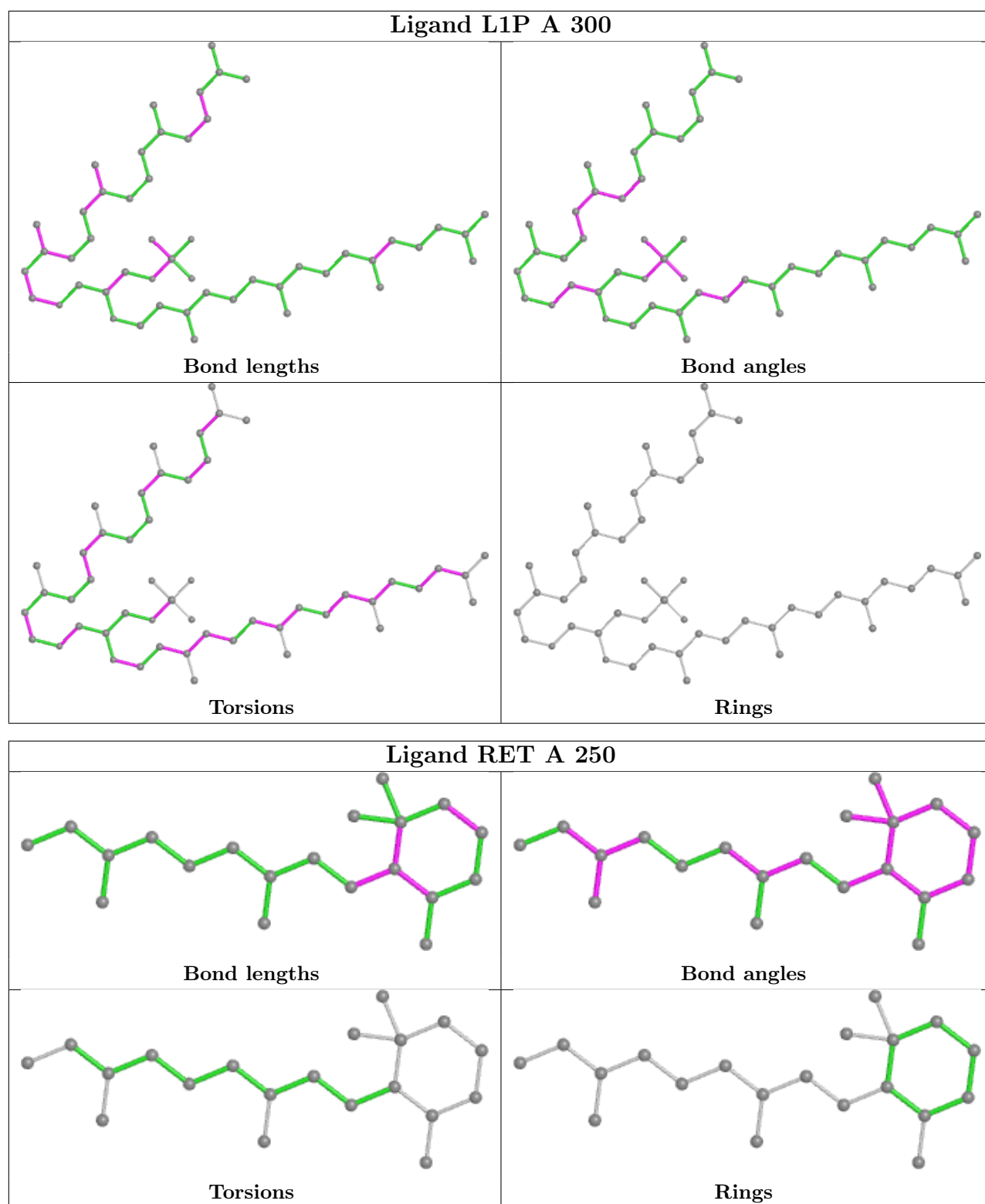
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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

There are no chain breaks in this entry.

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

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

Unable to reproduce the depositors R factor - this section is therefore empty.

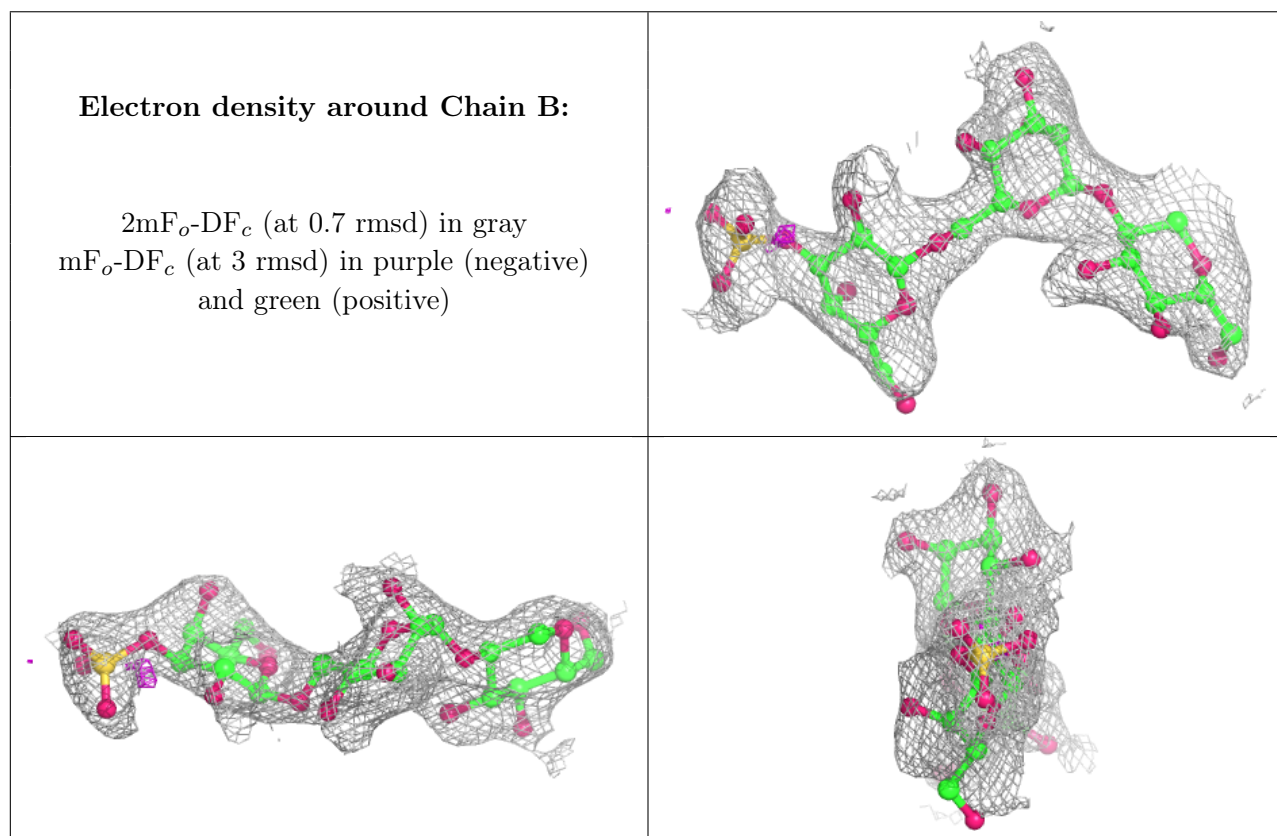
### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

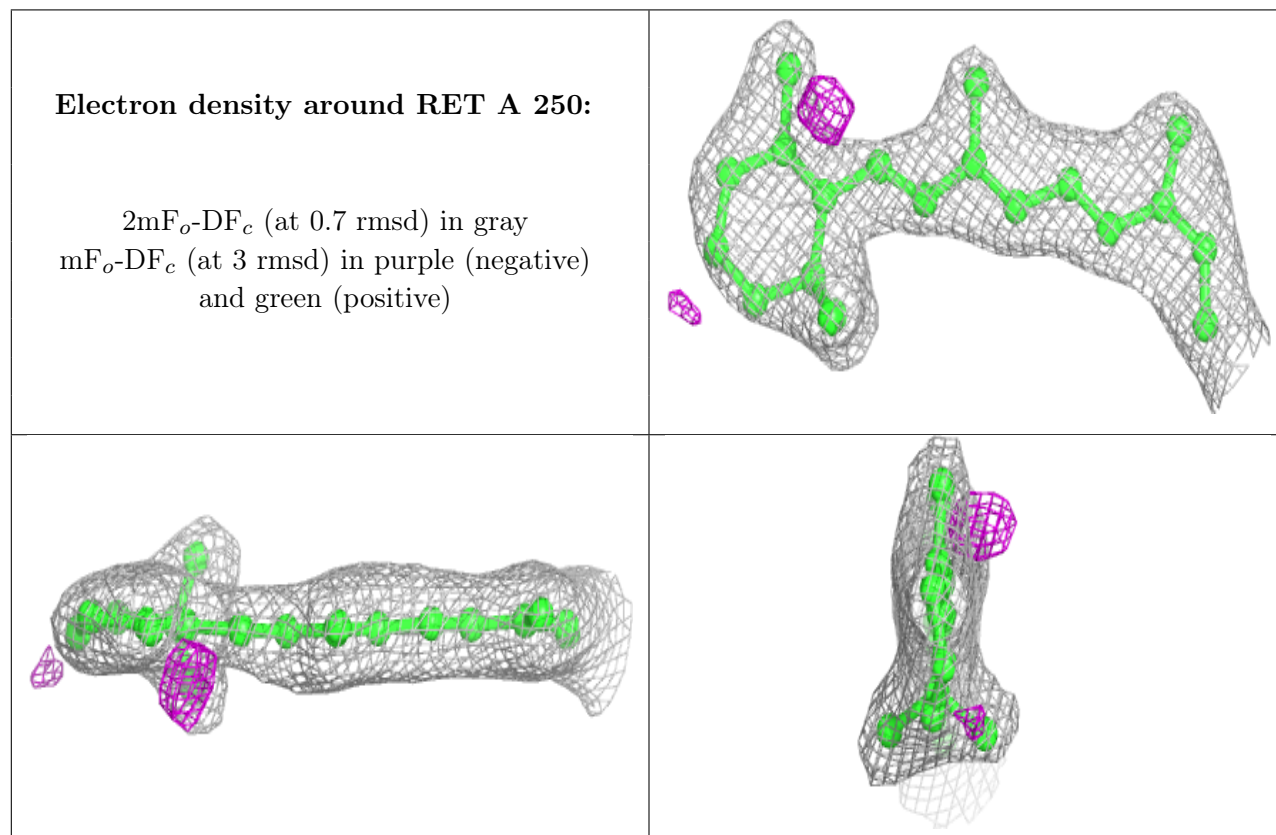
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



### 6.4 Ligands [i](#)

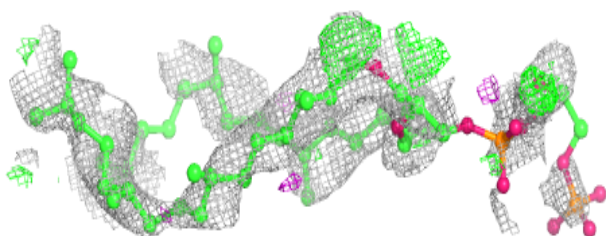
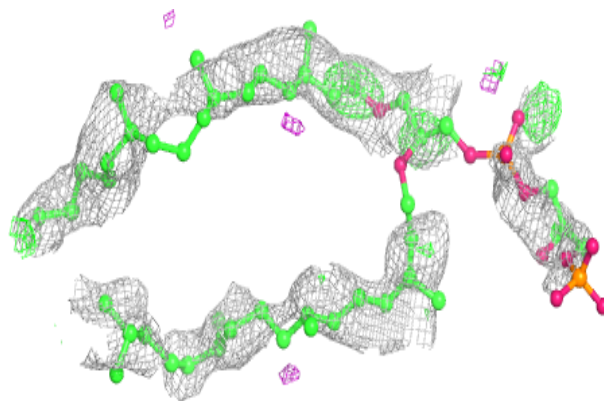
Unable to reproduce the depositors R factor - this section is therefore empty.

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.

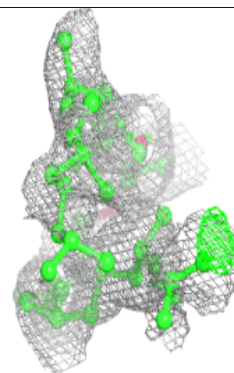
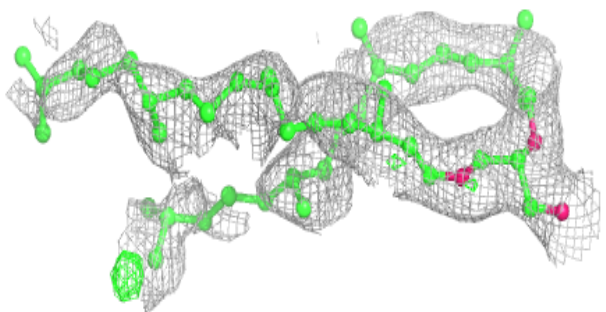
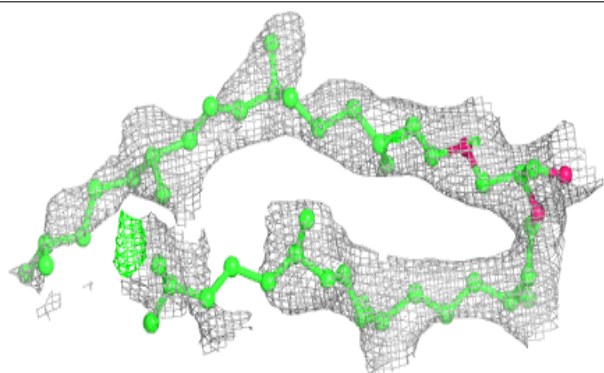


**Electron density around L3P A 260:**

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

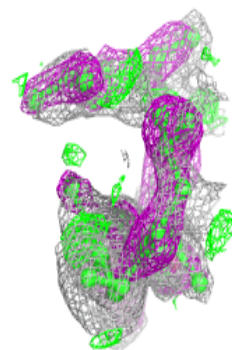
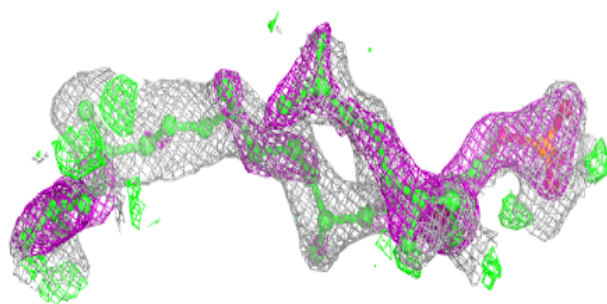
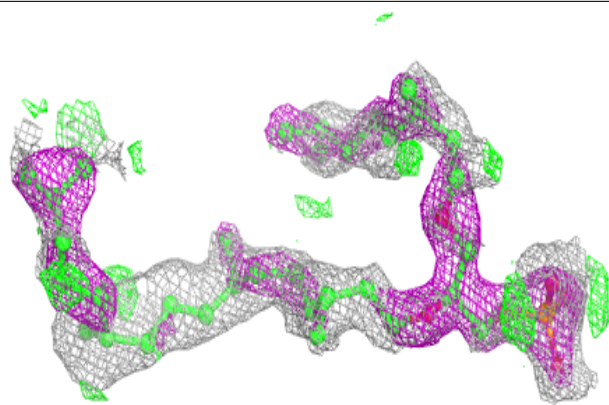
**Electron density around L2P A 270:**

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



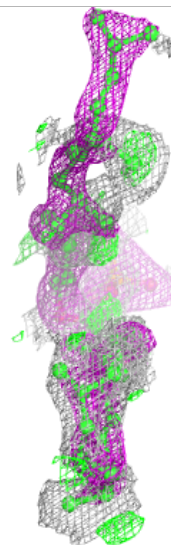
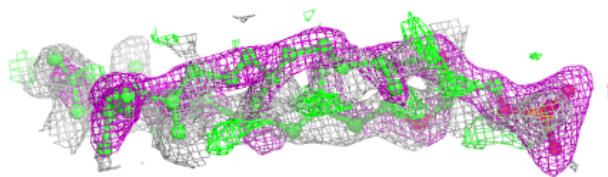
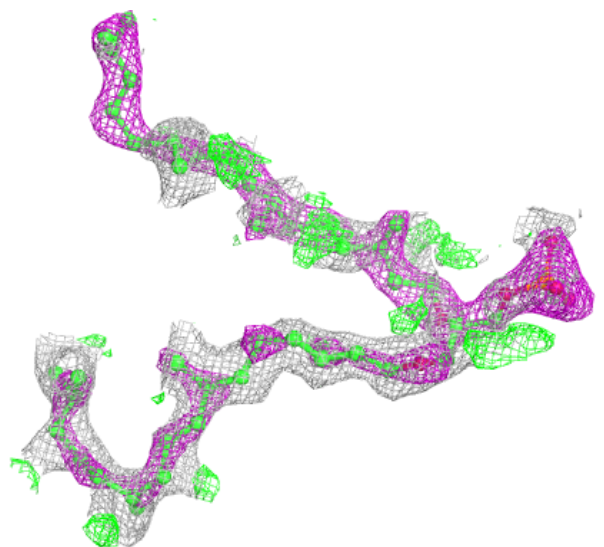
**Electron density around L1P A 280:**

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



**Electron density around L1P A 290:**

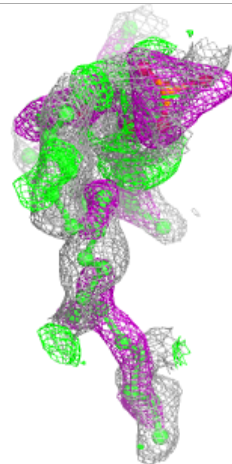
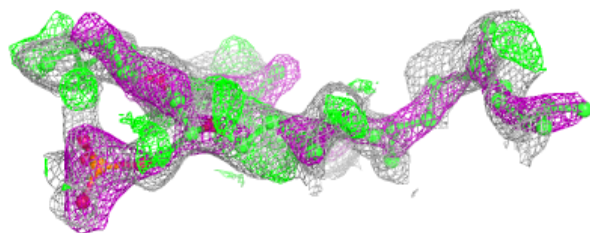
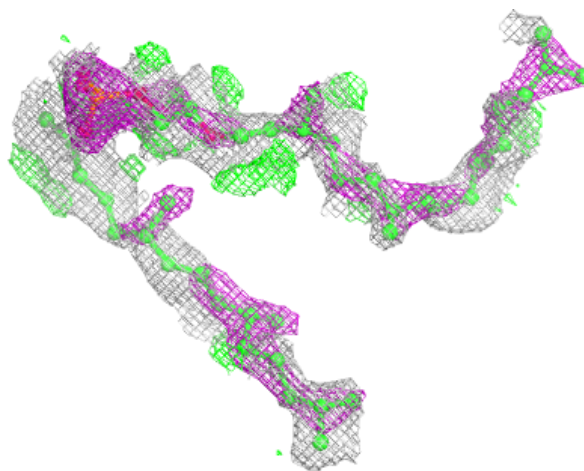
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

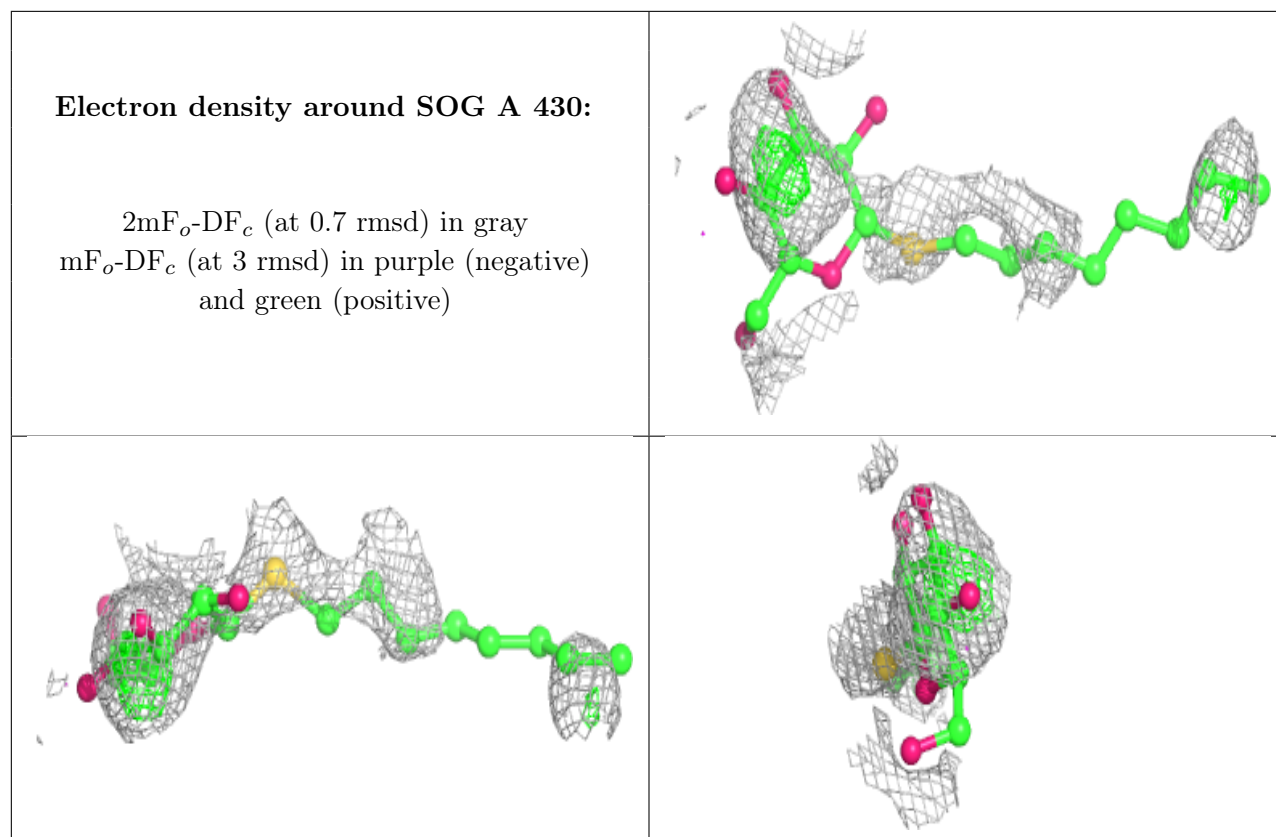




**Electron density around L1P A 300:**

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





## 6.5 Other polymers [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.