



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 10, 2023 – 06:34 pm GMT

PDB ID : 7YXA  
Title : XFEL crystal structure of the human sphingosine 1 phosphate receptor 5 in complex with ONO-5430608  
Authors : Lyapina, E.; Marin, E.; Gusach, A.; Orekhov, P.; Gerasimov, A.; Luginina, A.; Vakhrameev, D.; Ergasheva, M.; Kovaleva, M.; Khusainov, G.; Khorn, P.; Shevtsov, M.; Kovalev, K.; Okhrimenko, I.; Bukhdruker, S.; Popov, P.; Hu, H.; Weierstall, U.; Liu, W.; Cho, Y.; Gushchin, I.; Rogachev, A.; Bourenkov, G.; Park, S.; Park, G.; Huyn, H.J.; Park, J.; Gordeliy, V.; Borshchevskiy, V.; Mishin, A.; Cherezov, V.  
Deposited on : 2022-02-15  
Resolution : 2.20 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, 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

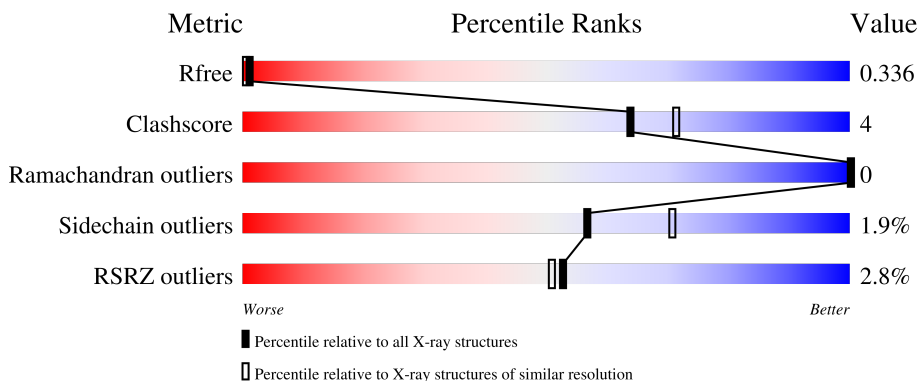
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	458	 2% 73% 9% 18%
1	B	458	 3% 72% 9% 19%

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

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12042 atoms, of which 6051 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 Sphingosine 1-phosphate receptor 5,Soluble cytochrome b562.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	375	5707	1815	2903	484	489	16	0	4	0
1	B	373	5727	1819	2919	480	492	17	0	4	0

There are 102 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-26	MET	-	initiating methionine	UNP Q9H228
A	-25	LYS	-	expression tag	UNP Q9H228
A	-24	THR	-	expression tag	UNP Q9H228
A	-23	ILE	-	expression tag	UNP Q9H228
A	-22	ILE	-	expression tag	UNP Q9H228
A	-21	ALA	-	expression tag	UNP Q9H228
A	-20	LEU	-	expression tag	UNP Q9H228
A	-19	SER	-	expression tag	UNP Q9H228
A	-18	TYR	-	expression tag	UNP Q9H228
A	-17	ILE	-	expression tag	UNP Q9H228
A	-16	PHE	-	expression tag	UNP Q9H228
A	-15	CYS	-	expression tag	UNP Q9H228
A	-14	LEU	-	expression tag	UNP Q9H228
A	-13	VAL	-	expression tag	UNP Q9H228
A	-12	PHE	-	expression tag	UNP Q9H228
A	-11	ALA	-	expression tag	UNP Q9H228
A	-10	ASP	-	expression tag	UNP Q9H228
A	-9	TYR	-	expression tag	UNP Q9H228
A	-8	LYS	-	expression tag	UNP Q9H228
A	-7	ASP	-	expression tag	UNP Q9H228
A	-6	ASP	-	expression tag	UNP Q9H228
A	-5	ASP	-	expression tag	UNP Q9H228
A	-4	ASP	-	expression tag	UNP Q9H228
A	-3	ALA	-	linker	UNP Q9H228
A	-2	GLY	-	linker	UNP Q9H228

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ARG	-	linker	UNP Q9H228
A	0	ALA	-	linker	UNP Q9H228
A	1007	TRP	MET	engineered mutation	UNP P0ABE7
A	1102	ILE	HIS	engineered mutation	UNP P0ABE7
A	1106	LEU	ARG	engineered mutation	UNP P0ABE7
A	322	GLY	-	linker	UNP Q9H228
A	323	ARG	-	linker	UNP Q9H228
A	324	PRO	-	linker	UNP Q9H228
A	325	LEU	-	expression tag	UNP Q9H228
A	326	GLU	-	expression tag	UNP Q9H228
A	327	VAL	-	expression tag	UNP Q9H228
A	328	LEU	-	expression tag	UNP Q9H228
A	329	PHE	-	expression tag	UNP Q9H228
A	330	GLN	-	expression tag	UNP Q9H228
A	331	GLY	-	expression tag	UNP Q9H228
A	332	PRO	-	expression tag	UNP Q9H228
A	333	HIS	-	expression tag	UNP Q9H228
A	334	HIS	-	expression tag	UNP Q9H228
A	335	HIS	-	expression tag	UNP Q9H228
A	336	HIS	-	expression tag	UNP Q9H228
A	337	HIS	-	expression tag	UNP Q9H228
A	338	HIS	-	expression tag	UNP Q9H228
A	339	HIS	-	expression tag	UNP Q9H228
A	340	HIS	-	expression tag	UNP Q9H228
A	341	HIS	-	expression tag	UNP Q9H228
A	342	HIS	-	expression tag	UNP Q9H228
B	-26	MET	-	initiating methionine	UNP Q9H228
B	-25	LYS	-	expression tag	UNP Q9H228
B	-24	THR	-	expression tag	UNP Q9H228
B	-23	ILE	-	expression tag	UNP Q9H228
B	-22	ILE	-	expression tag	UNP Q9H228
B	-21	ALA	-	expression tag	UNP Q9H228
B	-20	LEU	-	expression tag	UNP Q9H228
B	-19	SER	-	expression tag	UNP Q9H228
B	-18	TYR	-	expression tag	UNP Q9H228
B	-17	ILE	-	expression tag	UNP Q9H228
B	-16	PHE	-	expression tag	UNP Q9H228
B	-15	CYS	-	expression tag	UNP Q9H228
B	-14	LEU	-	expression tag	UNP Q9H228
B	-13	VAL	-	expression tag	UNP Q9H228
B	-12	PHE	-	expression tag	UNP Q9H228
B	-11	ALA	-	expression tag	UNP Q9H228

*Continued on next page...*

*Continued from previous page...*

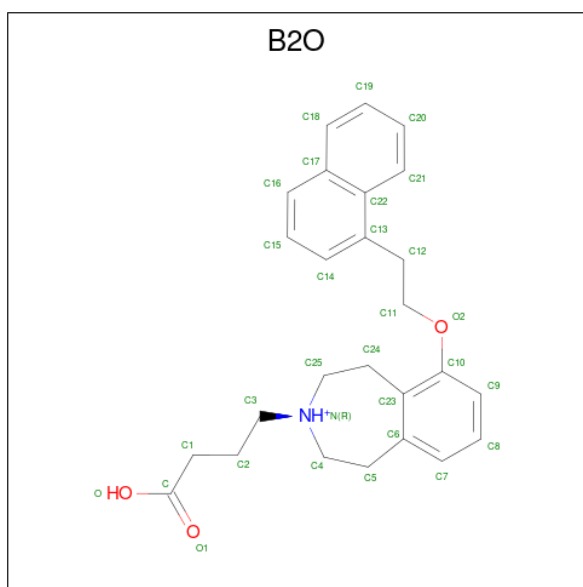
Chain	Residue	Modelled	Actual	Comment	Reference
B	-10	ASP	-	expression tag	UNP Q9H228
B	-9	TYR	-	expression tag	UNP Q9H228
B	-8	LYS	-	expression tag	UNP Q9H228
B	-7	ASP	-	expression tag	UNP Q9H228
B	-6	ASP	-	expression tag	UNP Q9H228
B	-5	ASP	-	expression tag	UNP Q9H228
B	-4	ASP	-	expression tag	UNP Q9H228
B	-3	ALA	-	linker	UNP Q9H228
B	-2	GLY	-	linker	UNP Q9H228
B	-1	ARG	-	linker	UNP Q9H228
B	0	ALA	-	linker	UNP Q9H228
B	1007	TRP	MET	engineered mutation	UNP P0ABE7
B	1102	ILE	HIS	engineered mutation	UNP P0ABE7
B	1106	LEU	ARG	engineered mutation	UNP P0ABE7
B	322	GLY	-	linker	UNP Q9H228
B	323	ARG	-	linker	UNP Q9H228
B	324	PRO	-	linker	UNP Q9H228
B	325	LEU	-	expression tag	UNP Q9H228
B	326	GLU	-	expression tag	UNP Q9H228
B	327	VAL	-	expression tag	UNP Q9H228
B	328	LEU	-	expression tag	UNP Q9H228
B	329	PHE	-	expression tag	UNP Q9H228
B	330	GLN	-	expression tag	UNP Q9H228
B	331	GLY	-	expression tag	UNP Q9H228
B	332	PRO	-	expression tag	UNP Q9H228
B	333	HIS	-	expression tag	UNP Q9H228
B	334	HIS	-	expression tag	UNP Q9H228
B	335	HIS	-	expression tag	UNP Q9H228
B	336	HIS	-	expression tag	UNP Q9H228
B	337	HIS	-	expression tag	UNP Q9H228
B	338	HIS	-	expression tag	UNP Q9H228
B	339	HIS	-	expression tag	UNP Q9H228
B	340	HIS	-	expression tag	UNP Q9H228
B	341	HIS	-	expression tag	UNP Q9H228
B	342	HIS	-	expression tag	UNP Q9H228

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
2	A	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
2	B	1	Total	C	H	N	O	0	0
			28	8	14	1	5		

- Molecule 3 is 4-[6-(2-naphthalen-1-ylethoxy)-2,3,4,5-tetrahydro-1H-3-benzazepin-3-ium-3-yl]butanoic acid (three-letter code: B2O) (formula: C<sub>26</sub>H<sub>30</sub>NO<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



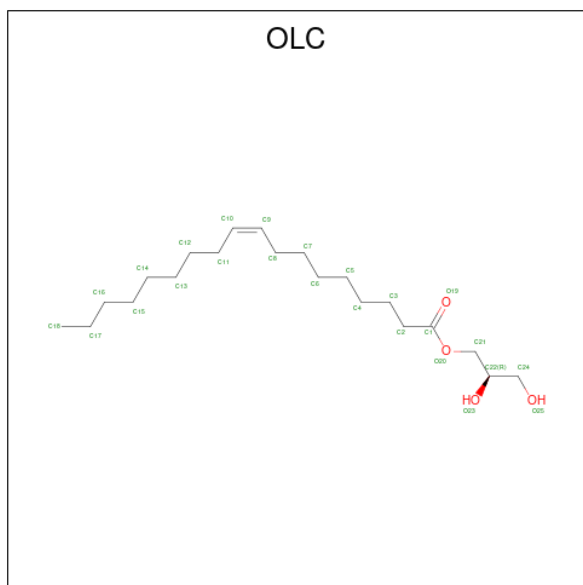
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
3	A	1	Total	C	H	N	O	0	0
			59	26	29	1	3		

*Continued on next page...*

Continued from previous page...

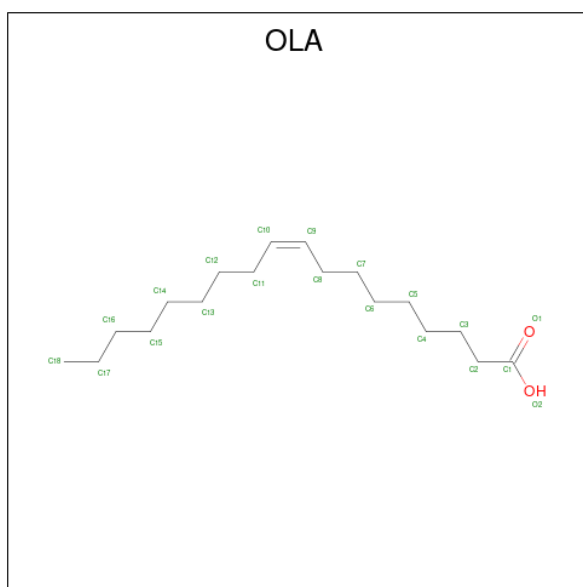
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	H	N			O
3	B	1	59	26	29	1	3	0	0

- Molecule 4 is (2R)-2,3-dihydroxypropyl (9Z)-octadec-9-enoate (three-letter code: OLC) (formula: C<sub>21</sub>H<sub>40</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
4	A	1	55	18	33	4	0	0
4	A	1	52	17	31	4	0	0
4	A	1	65	21	40	4	0	0

- Molecule 5 is OLEIC ACID (three-letter code: OLA) (formula: C<sub>18</sub>H<sub>34</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	H	O	0	0
			53	18	33	2		
5	B	1	Total	C	H		0	0
			15	9	6			

- Molecule 6 is water.

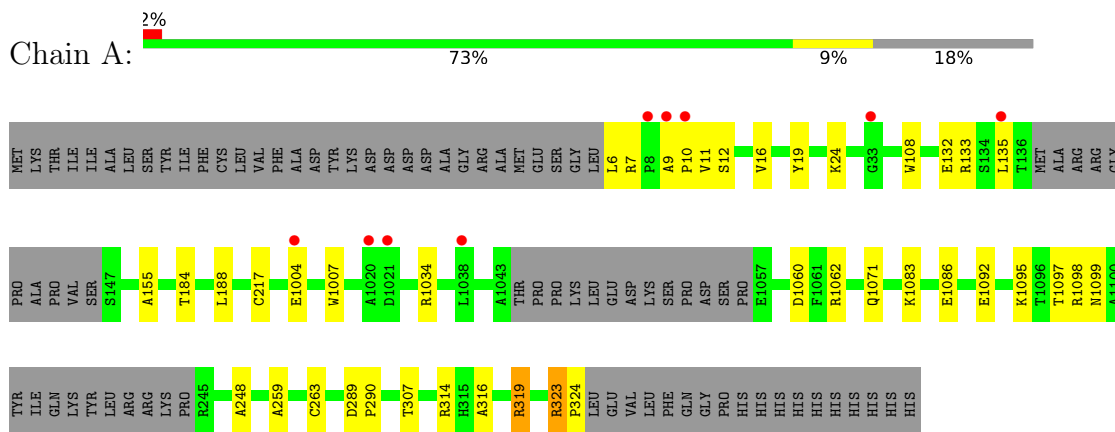
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	111	Total	O	0	0
			111	111		
6	B	83	Total	O	0	0
			83	83		



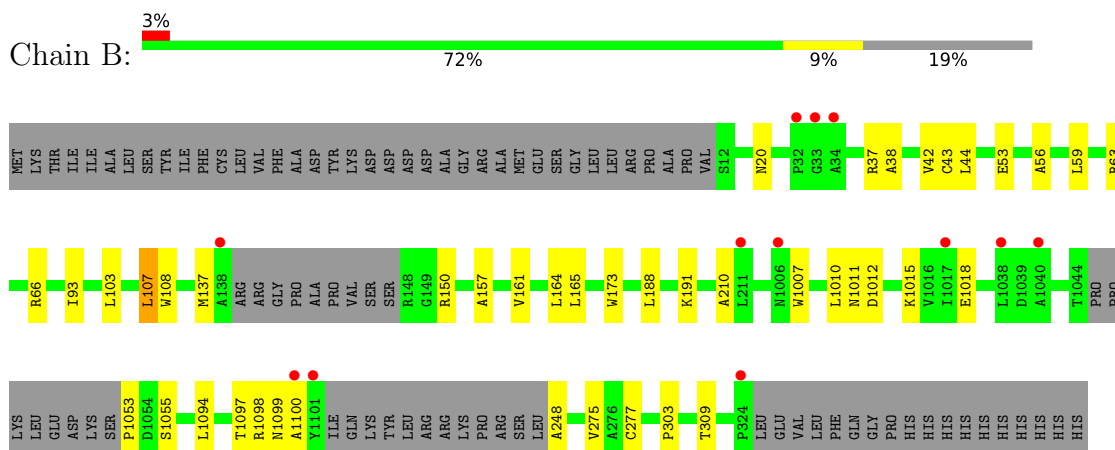
### 3 Residue-property plots [i](#)

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: Spingosine 1-phosphate receptor 5,Soluble cytochrome b562



- Molecule 1: Spingosine 1-phosphate receptor 5,Soluble cytochrome b562



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.84Å 103.42Å 187.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.97 – 2.20 29.97 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.97-2.20) 100.0 (29.97-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.43 (at 2.00Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.301 , 0.329 0.307 , 0.336	Depositor DCC
$R_{free}$ test set	2650 reflections (3.33%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.7	Xtrriage
Anisotropy	0.102	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 111.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.24$ , $\langle L^2 \rangle = 0.10$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	12042	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 30.68 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2559e-03.*

<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: OLC, B2O, OLA, 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	0.25	0/2865	0.42	0/3908
1	B	0.24	0/2872	0.40	0/3915
All	All	0.24	0/5737	0.41	0/7823

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	2804	2903	2885	25	0
1	B	2808	2919	2903	24	0
2	A	14	14	13	0	0
2	B	14	14	13	0	0
3	A	30	29	0	1	0
3	B	30	29	0	0	0
4	A	68	104	100	6	0
5	B	29	39	45	2	0
6	A	111	0	0	6	0
6	B	83	0	0	2	0
All	All	5991	6051	5959	50	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 4.

All (50) 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:11:VAL:HG23	1:A:12:SER:H	1.55	0.71
1:A:307:THR:O	1:A:314:ARG:NH1	2.25	0.69
1:A:6:LEU:O	1:A:7:ARG:NH1	2.28	0.66
1:A:314:ARG:NH2	6:A:2604:HOH:O	2.31	0.61
1:B:191:LYS:NZ	1:B:277:CYS:O	2.33	0.60
4:A:2505:OLC:H14A	1:B:161:VAL:HG22	1.83	0.60
1:A:1062:ARG:NH2	6:A:2608:HOH:O	2.38	0.56
1:B:1094:LEU:HA	1:B:1097:THR:HG22	1.90	0.54
1:A:7:ARG:NH2	6:A:2609:HOH:O	2.39	0.54
1:A:24:LYS:HD2	1:A:184:THR:O	2.10	0.51
1:A:19:TYR:OH	3:A:2502:B2O:O	2.25	0.51
1:A:133:ARG:NH1	1:A:248:ALA:O	2.44	0.51
1:B:1012:ASP:OD1	6:B:2601:HOH:O	2.20	0.50
1:A:108:TRP:CZ3	1:A:188:LEU:HB2	2.47	0.49
1:A:1071:GLN:OE1	1:A:1097:THR:HG21	2.13	0.49
4:A:2505:OLC:H21A	1:B:173:TRP:HA	1.94	0.48
1:A:9:ALA:HB1	1:A:10:PRO:HD2	1.96	0.48
1:A:323:ARG:HG3	1:A:324:PRO:HD3	1.94	0.48
1:B:210:ALA:HA	5:B:2503:OLA:H31	1.96	0.48
1:B:108:TRP:CZ3	1:B:188:LEU:HB2	2.49	0.47
4:A:2504:OLC:H12	1:B:164:LEU:HD13	1.96	0.47
1:B:38:ALA:O	1:B:42:VAL:HG23	2.15	0.47
1:B:191:LYS:HB3	1:B:275:VAL:HG13	1.98	0.46
1:B:157:ALA:O	1:B:161:VAL:HG23	2.16	0.45
1:A:259:ALA:O	1:A:263:CYS:N	2.44	0.45
1:B:1099:ASN:OD1	1:B:1100:ALA:N	2.48	0.45
1:B:1011:ASN:OD1	1:B:1098:ARG:NH1	2.49	0.45
1:A:1092:GLU:OE1	1:A:1095:LYS:HD3	2.16	0.45
1:A:1004:GLU:HA	1:A:1007:TRP:HB3	1.99	0.45
1:A:1007:TRP:CH2	1:A:1099:ASN:HA	2.52	0.45
5:B:2503:OLA:H9	5:B:2504:OLA:H122	1.99	0.45
1:B:53:GLU:HG3	1:B:303:PRO:HG2	1.99	0.44
4:A:2505:OLC:H11A	1:B:165:LEU:HD13	1.99	0.44
1:B:1053:PRO:O	1:B:1055:SER:N	2.48	0.44
1:A:290:PRO:HG3	4:A:2503:OLC:H4A	1.99	0.44
1:B:1015:LYS:O	1:B:1018:GLU:HG2	2.19	0.43
1:B:20:ASN:C	1:B:20:ASN:OD1	2.57	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:248:ALA:HA	1:B:309:THR:O	2.19	0.42
1:A:1098:ARG:NH2	6:A:2626:HOH:O	2.49	0.42
1:B:44:LEU:HD21	1:B:93:ILE:HB	2.02	0.41
1:B:56:ALA:HA	1:B:59:LEU:HD12	2.02	0.41
1:A:316:ALA:O	1:A:319:ARG:NH1	2.53	0.41
1:A:16:VAL:HG12	6:A:2672:HOH:O	2.21	0.41
1:A:1034:ARG:NH1	6:A:2631:HOH:O	2.53	0.41
1:A:1083:LYS:HD2	1:A:1086:GLU:HG3	2.03	0.41
1:B:63:ARG:NH1	6:B:2609:HOH:O	2.48	0.41
1:A:155:ALA:HB1	4:A:2504:OLC:H6	2.02	0.41
1:A:132:GLU:O	1:A:135:LEU:HG	2.22	0.40
1:B:103:LEU:HD22	1:B:107:LEU:HD13	2.04	0.40
1:B:1007:TRP:HA	1:B:1010:LEU:HD12	2.04	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	371/458 (81%)	359 (97%)	12 (3%)	0	100	100
1	B	369/458 (81%)	360 (98%)	9 (2%)	0	100	100
All	All	740/916 (81%)	719 (97%)	21 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	285/364 (78%)	280 (98%)	5 (2%)	59	72
1	B	288/364 (79%)	282 (98%)	6 (2%)	53	67
All	All	573/728 (79%)	562 (98%)	11 (2%)	57	71

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	217	CYS
1	A	1060	ASP
1	A	289	ASP
1	A	319	ARG
1	A	323	ARG
1	B	37	ARG
1	B	43	CYS
1	B	66	ARG
1	B	107	LEU
1	B	137	MET
1	B	150	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	68	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](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

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
2	NAG	A	2501	1	14,14,15	0.52	0	17,19,21	0.63	1 (5%)
3	B2O	A	2502	-	33,33,33	0.62	0	40,44,44	0.46	0
4	OLC	A	2504	-	20,20,24	0.77	0	21,21,25	0.94	0
5	OLA	B	2503	-	19,19,19	0.56	0	19,19,19	0.86	1 (5%)
4	OLC	A	2503	-	21,21,24	0.63	0	22,22,25	0.93	0
5	OLA	B	2504	-	8,8,19	0.30	0	7,7,19	0.28	0
2	NAG	B	2501	1	14,14,15	0.62	0	17,19,21	0.85	1 (5%)
3	B2O	B	2502	-	33,33,33	0.62	0	40,44,44	0.56	1 (2%)
4	OLC	A	2505	-	24,24,24	0.67	0	25,25,25	0.84	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
2	NAG	A	2501	1	-	0/6/23/26	0/1/1/1
3	B2O	A	2502	-	-	6/12/22/22	0/4/4/4
4	OLC	A	2504	-	-	8/20/20/24	-
5	OLA	B	2503	-	-	10/17/17/17	-
4	OLC	A	2503	-	-	8/21/21/24	-
5	OLA	B	2504	-	-	4/6/6/17	-
2	NAG	B	2501	1	-	2/6/23/26	0/1/1/1
3	B2O	B	2502	-	-	0/12/22/22	0/4/4/4
4	OLC	A	2505	-	-	10/24/24/24	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2501	NAG	C1-O5-C5	2.80	115.98	112.19
3	B	2502	B2O	C5-C4-N	-2.45	109.75	115.04
2	A	2501	NAG	C1-O5-C5	2.18	115.14	112.19
5	B	2503	OLA	O2-C1-C2	2.07	120.68	114.03

There are no chirality outliers.

All (48) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	2502	B2O	C-C1-C2-C3
5	B	2504	OLA	C10-C11-C12-C13
4	A	2504	OLC	O20-C21-C22-O23
4	A	2503	OLC	C2-C1-O20-C21
4	A	2503	OLC	O19-C1-O20-C21
3	A	2502	B2O	C1-C2-C3-N
4	A	2504	OLC	O20-C21-C22-C24
2	B	2501	NAG	C8-C7-N2-C2
2	B	2501	NAG	O7-C7-N2-C2
5	B	2503	OLA	C12-C13-C14-C15
5	B	2503	OLA	C2-C3-C4-C5
4	A	2505	OLC	C12-C13-C14-C15
5	B	2503	OLA	C11-C12-C13-C14
4	A	2505	OLC	C21-C22-C24-O25
4	A	2504	OLC	C10-C11-C12-C13
4	A	2504	OLC	C5-C6-C7-C8
5	B	2503	OLA	C5-C6-C7-C8
4	A	2505	OLC	O23-C22-C24-O25
4	A	2505	OLC	C2-C1-O20-C21
4	A	2505	OLC	O19-C1-O20-C21
4	A	2504	OLC	C6-C7-C8-C9
4	A	2505	OLC	C4-C5-C6-C7
4	A	2503	OLC	C5-C6-C7-C8
5	B	2503	OLA	C15-C16-C17-C18
4	A	2505	OLC	C13-C14-C15-C16
4	A	2503	OLC	C10-C11-C12-C13
4	A	2503	OLC	C1-C2-C3-C4
5	B	2504	OLA	C9-C10-C11-C12
5	B	2503	OLA	C4-C5-C6-C7
4	A	2504	OLC	C1-C2-C3-C4
5	B	2503	OLA	C10-C11-C12-C13
4	A	2505	OLC	C15-C16-C17-C18
5	B	2503	OLA	O2-C1-C2-C3
5	B	2503	OLA	O1-C1-C2-C3

*Continued on next page...*



*Continued from previous page...*

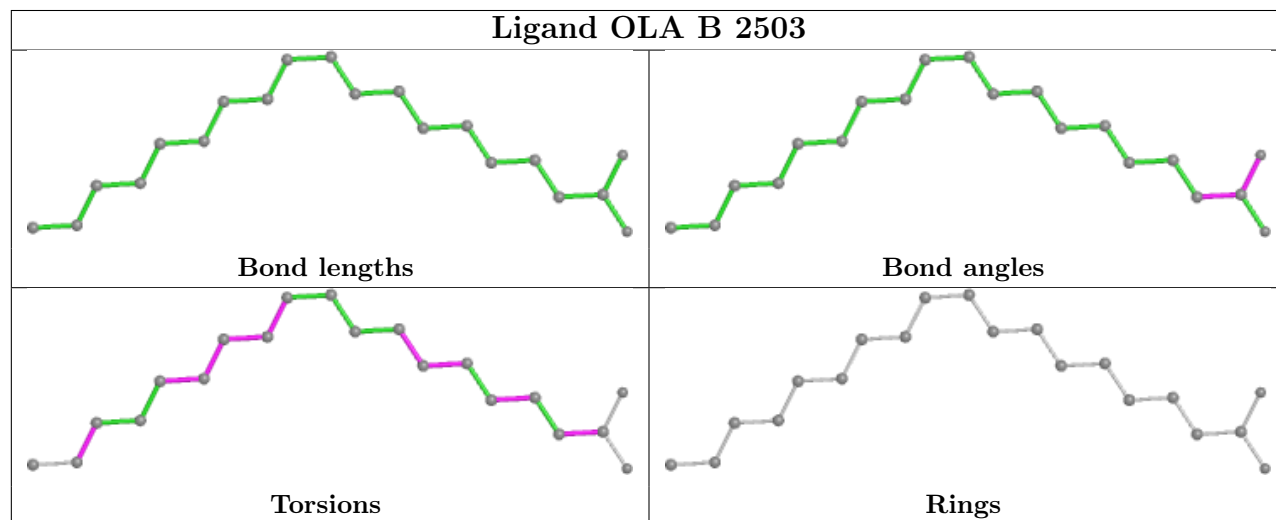
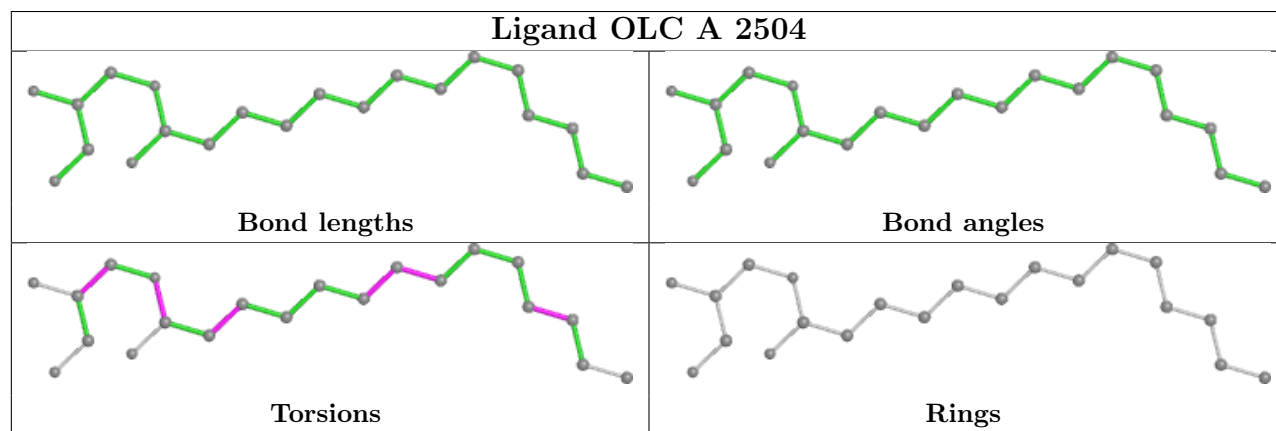
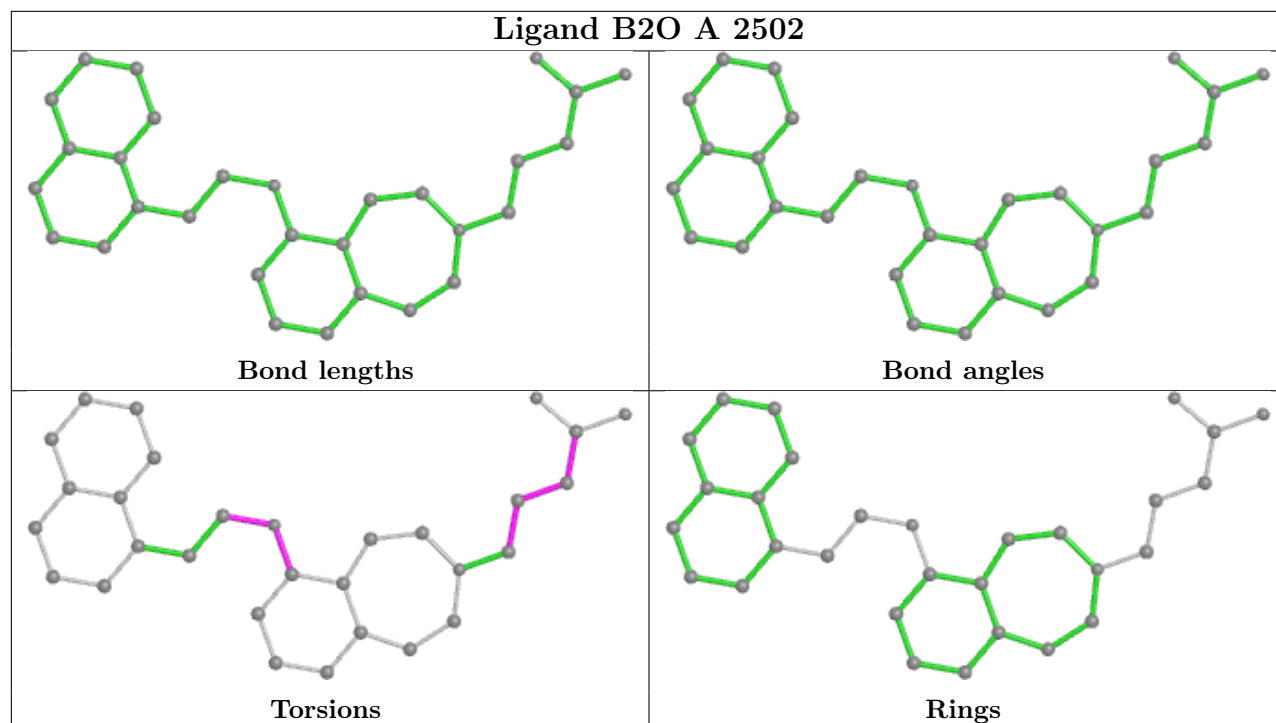
Mol	Chain	Res	Type	Atoms
4	A	2503	OLC	C12-C13-C14-C15
3	A	2502	B2O	O1-C-C1-C2
3	A	2502	B2O	O-C-C1-C2
5	B	2504	OLA	C5-C6-C7-C8
4	A	2504	OLC	C2-C1-O20-C21
4	A	2505	OLC	C14-C15-C16-C17
3	A	2502	B2O	C12-C11-O2-C10
4	A	2504	OLC	O19-C1-O20-C21
5	B	2503	OLA	C9-C10-C11-C12
5	B	2504	OLA	C6-C7-C8-C9
4	A	2503	OLC	O20-C1-C2-C3
4	A	2503	OLC	C4-C5-C6-C7
3	A	2502	B2O	C9-C10-O2-C11
4	A	2505	OLC	C7-C8-C9-C10

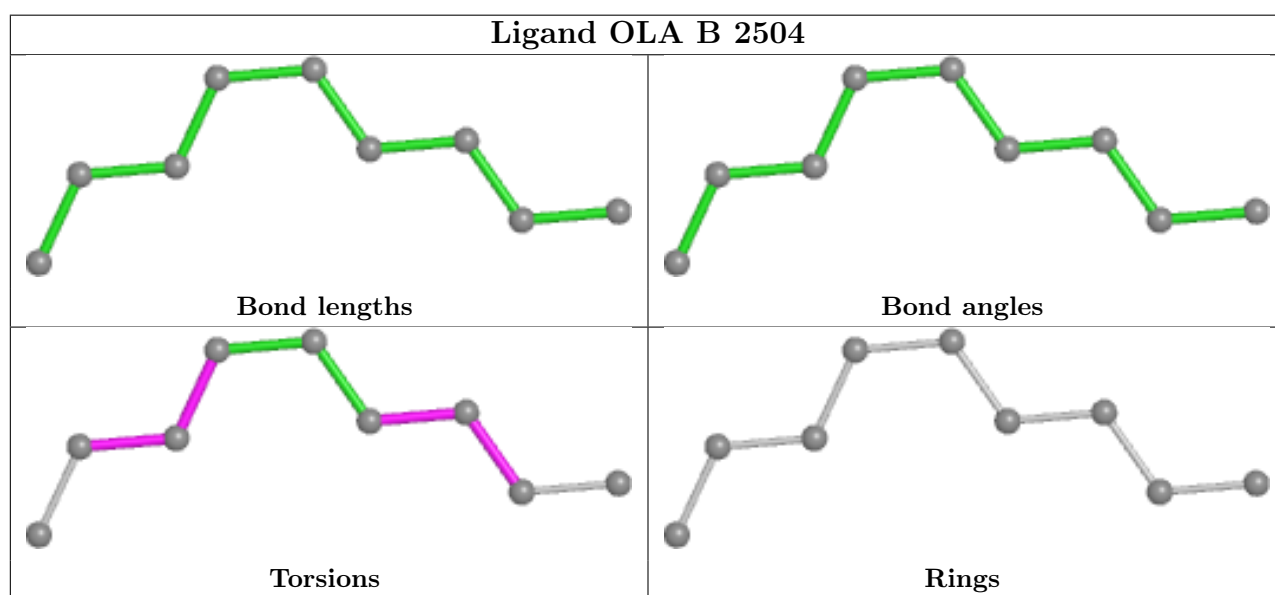
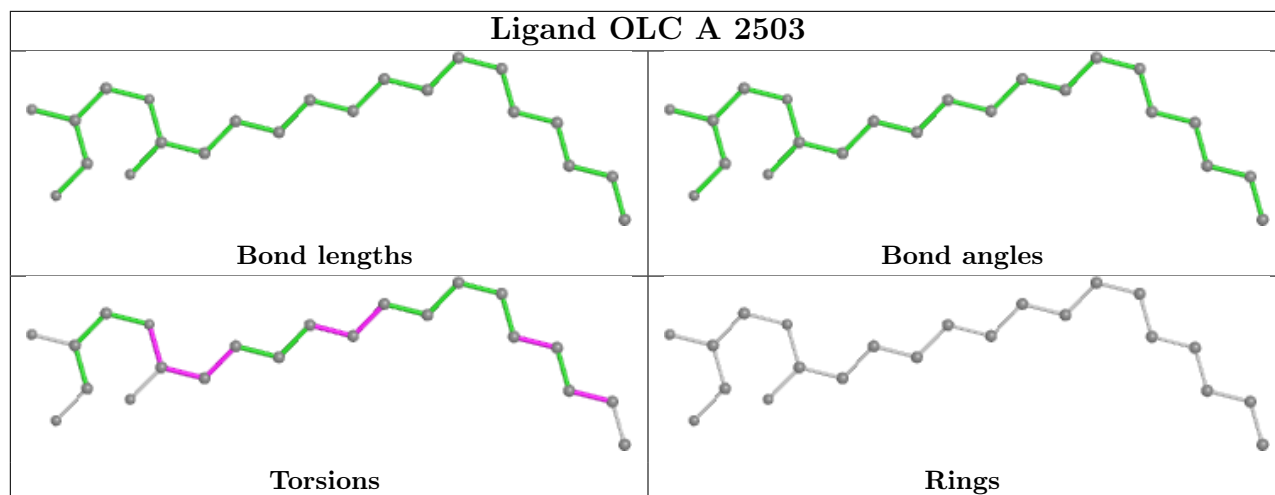
There are no ring outliers.

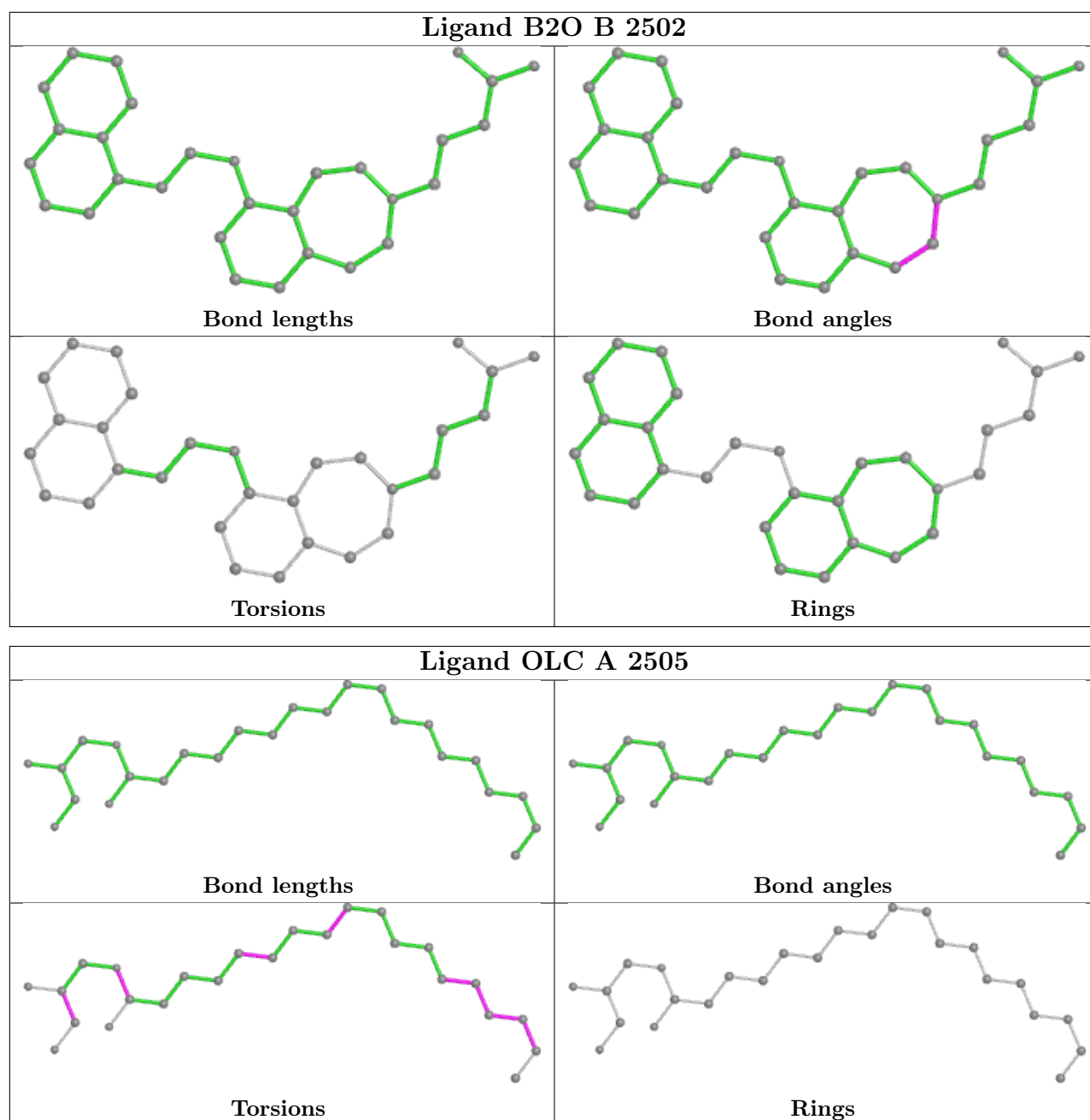
6 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2502	B2O	1	0
4	A	2504	OLC	2	0
5	B	2503	OLA	2	0
4	A	2503	OLC	1	0
5	B	2504	OLA	1	0
4	A	2505	OLC	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

### 6.1 Protein, DNA and RNA chains

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	375/458 (81%)	-0.03	9 (2%) 59 56	20, 37, 64, 79	0
1	B	373/458 (81%)	0.08	12 (3%) 47 45	33, 46, 65, 91	0
All	All	748/916 (81%)	0.02	21 (2%) 53 51	20, 44, 65, 91	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	138	ALA	5.4
1	B	33	GLY	4.7
1	A	1038	LEU	4.4
1	A	33	GLY	3.4
1	B	1038	LEU	3.4
1	B	34	ALA	3.0
1	A	8	PRO	2.9
1	B	211	LEU	2.9
1	B	1100	ALA	2.8
1	A	9	ALA	2.8
1	B	32	PRO	2.7
1	B	324	PRO	2.7
1	B	1101	TYR	2.6
1	A	135	LEU	2.5
1	B	1006	ASN	2.3
1	A	1020	ALA	2.3
1	A	1004	GLU	2.1
1	B	1017	ILE	2.1
1	A	1021	ASP	2.0
1	B	1040	ALA	2.0
1	A	10	PRO	2.0

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

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

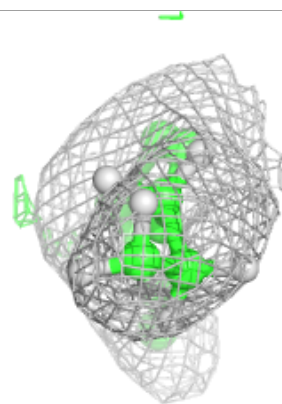
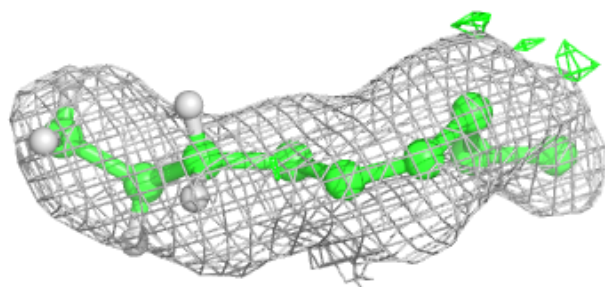
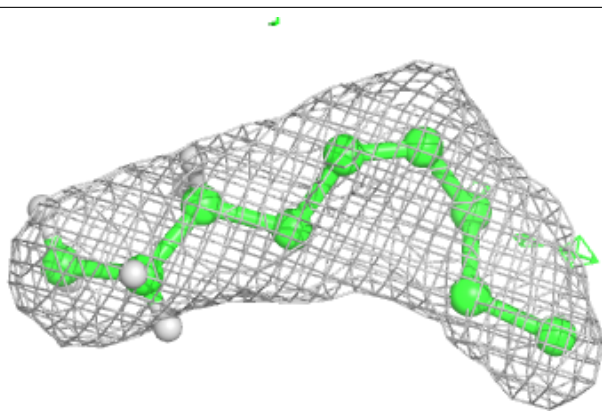
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	A	2501	14/15	0.82	0.14	46,70,88,91	0
2	NAG	B	2501	14/15	0.84	0.13	33,66,84,92	0
5	OLA	B	2504	9/20	0.84	0.26	35,57,69,69	0
4	OLC	A	2503	22/25	0.85	0.21	12,56,81,90	0
5	OLA	B	2503	20/20	0.86	0.17	26,51,72,77	0
3	B2O	B	2502	30/30	0.89	0.16	9,31,54,65	0
3	B2O	A	2502	30/30	0.90	0.15	11,37,65,78	0
4	OLC	A	2504	21/25	0.92	0.15	9,42,76,92	0
4	OLC	A	2505	25/25	0.94	0.13	8,49,76,87	0

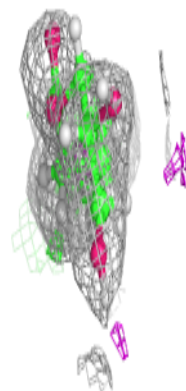
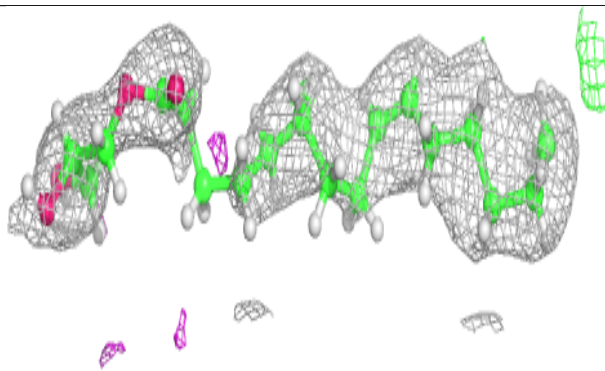
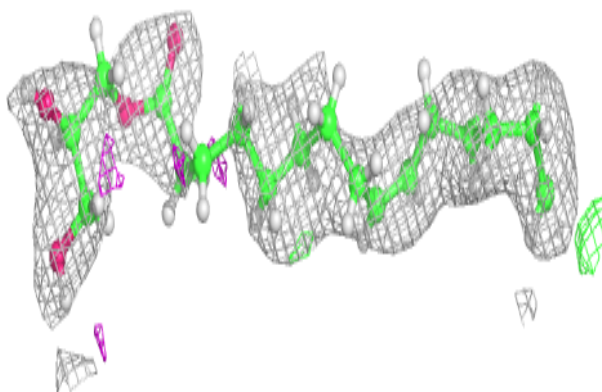
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

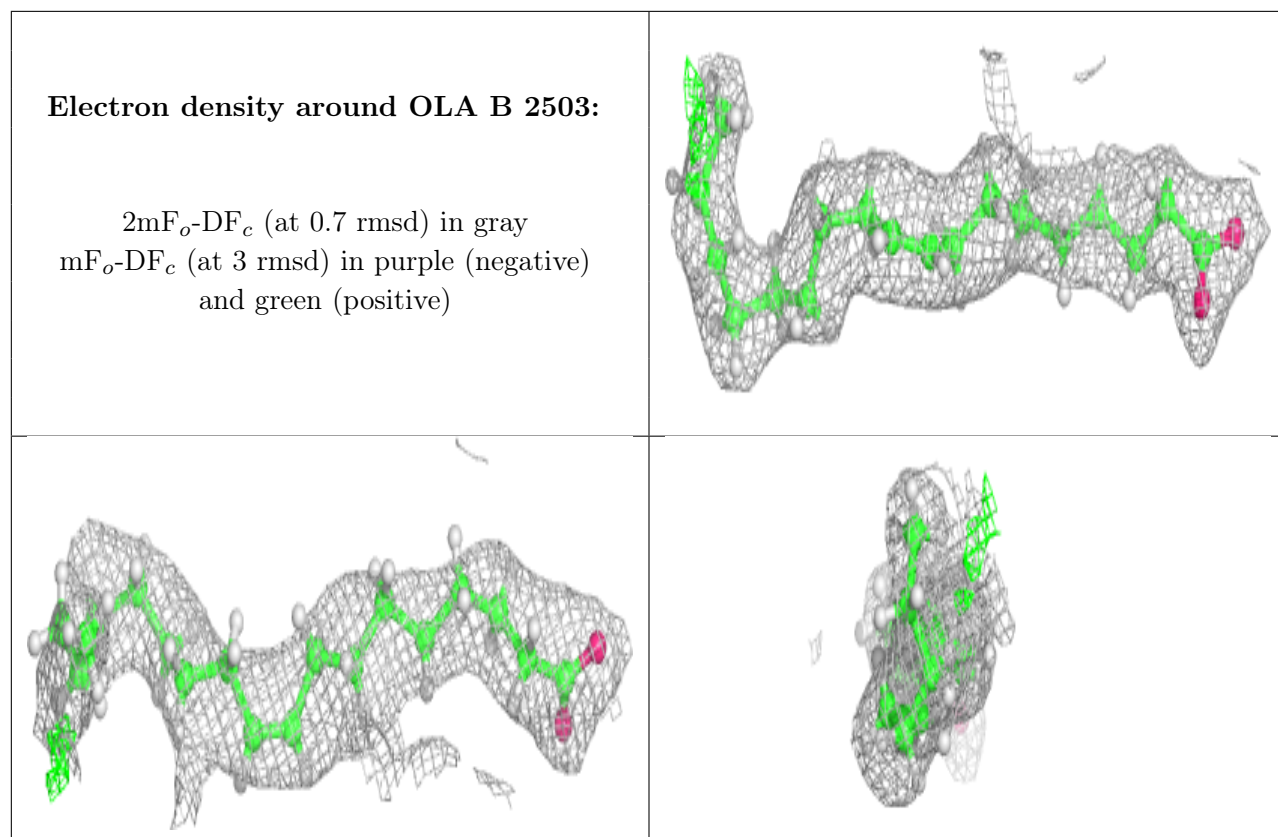
**Electron density around OLA B 2504:**

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

**Electron density around OLC A 2503:**

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

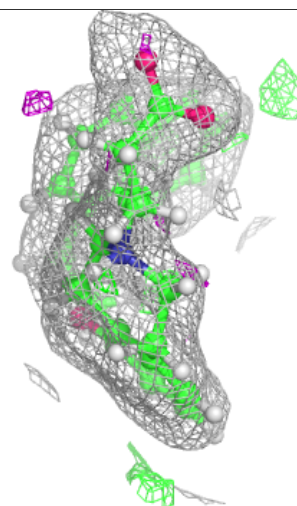
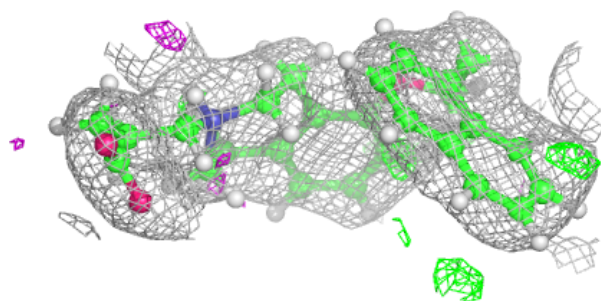
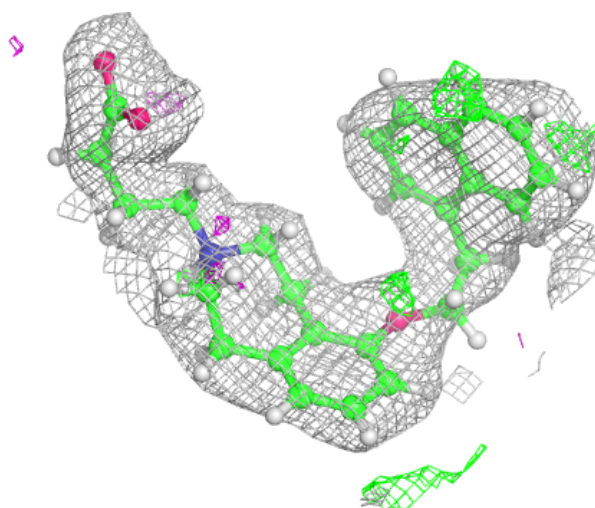






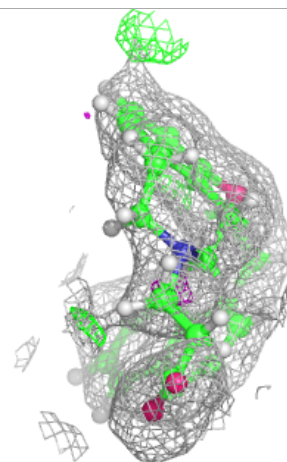
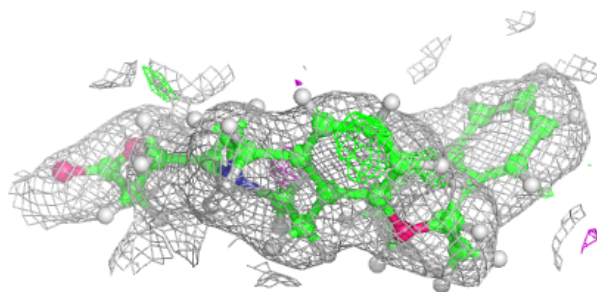
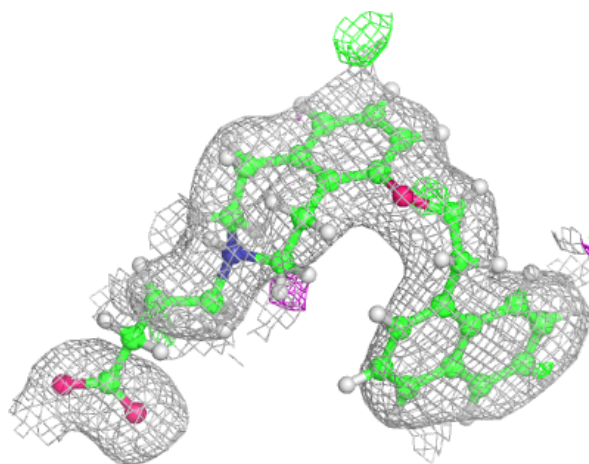
**Electron density around B2O B 2502:**

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



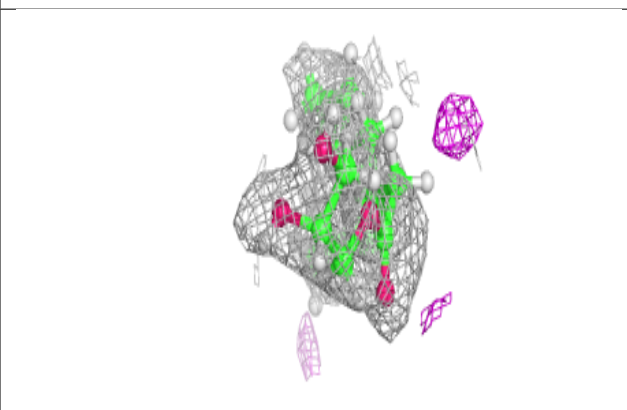
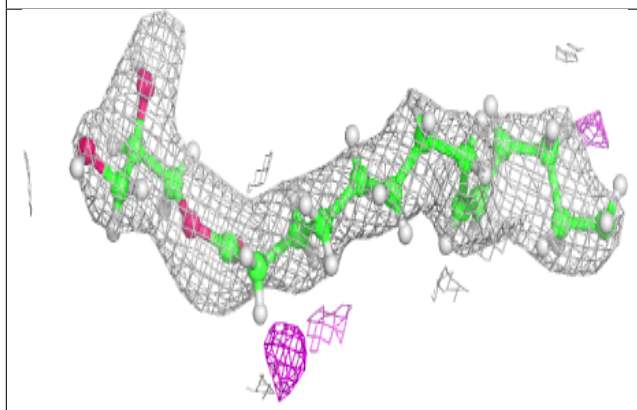
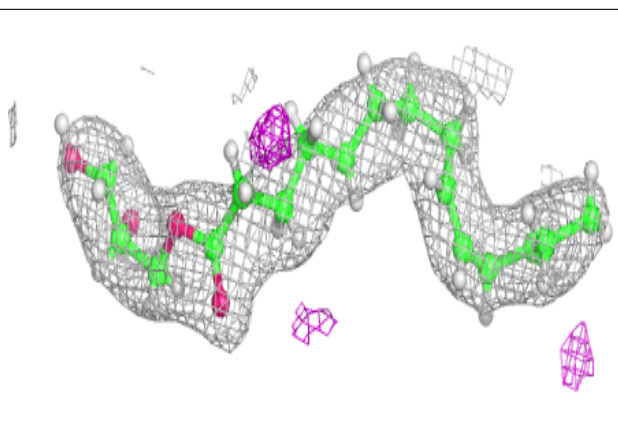
**Electron density around B2O A 2502:**

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

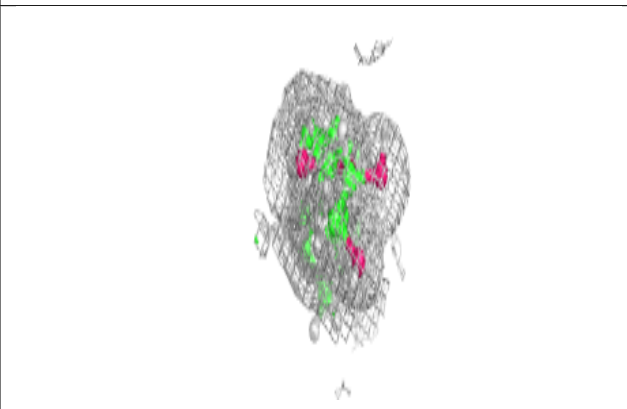
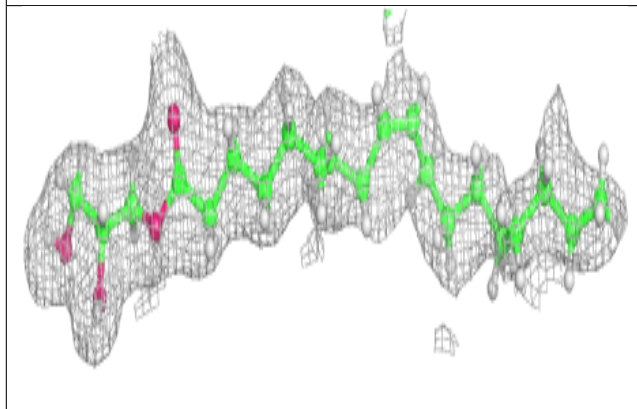
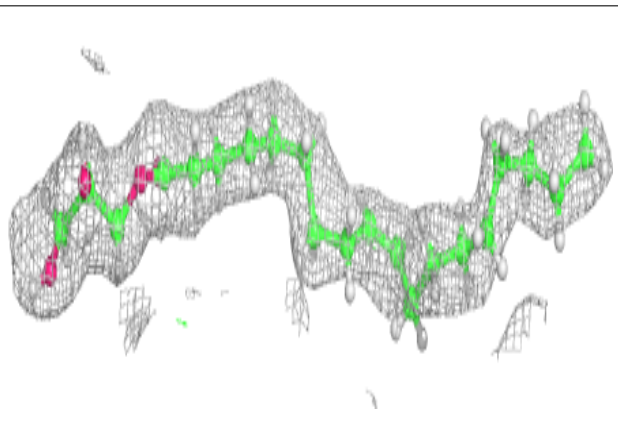


**Electron density around OLC A 2504:**

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

**Electron density around OLC A 2505:**

$2mF_o-DF_c$  (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.