



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 3, 2022 – 12:20 pm GMT

PDB ID : 7O8H
Title : NmHR light state structure at 10 ns after photoexcitation determined by serial femtosecond crystallography (with extrapolated, dark and light dataset)
Authors : Mous, S.; Gotthard, G.; Ehrenberg, D.; Sen, S.; James, D.; Johnson, P.; Weinert, T.; Nass, K.; Furrer, A.; Kekilli, D.; Ma, P.; Bruenle, S.; Casadei, C.; Martiel, I.; Dworkowski, F.; Gashi, D.; Skopintsev, P.; Wranik, M.; Knopp, G.; Panepucci, E.; Panneels, V.; Cirelli, C.; Ozerov, D.; Schertler, G.; Wang, M.; Milne, C.; Standfuss, J.; Schapiro, I.; Heberle, J.; Nogly, P.
Deposited on : 2021-04-15
Resolution : 1.80 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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

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)
Xtriage (Phenix) : 1.13
EDS : 2.26
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.26

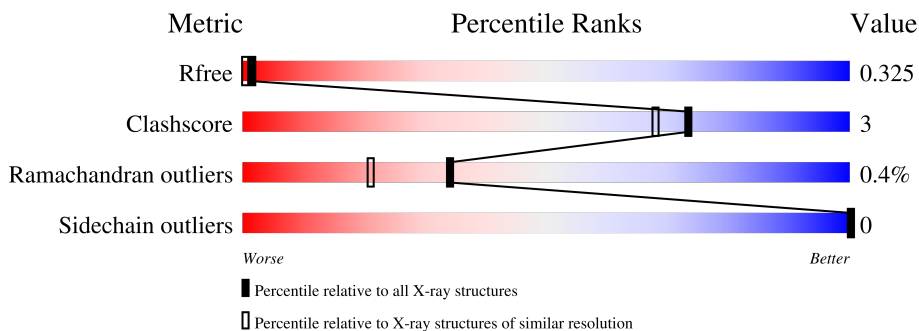
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)

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 ≥ 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	296	 83% 6% 11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CL	A	302	-	-	X	-

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 2290 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.

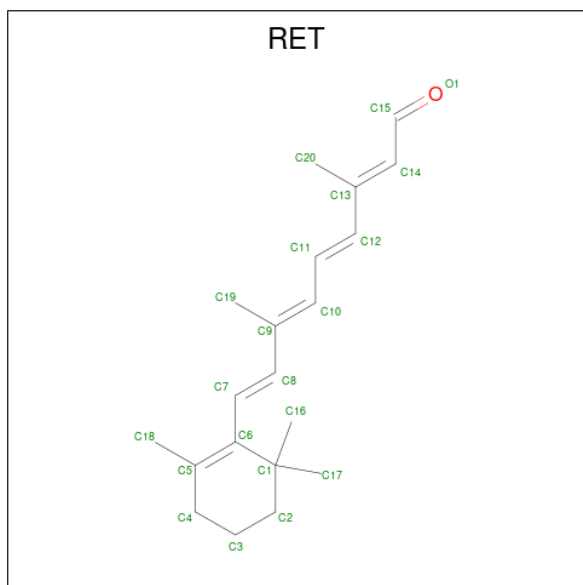
- Molecule 1 is a protein called Chloride pumping rhodopsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	262	2084	1391	317	361	15	0	5	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-15	MET	-	initiating methionine	UNP W8VZW3
A	-14	ALA	-	expression tag	UNP W8VZW3
A	-13	SER	-	expression tag	UNP W8VZW3
A	-12	MET	-	expression tag	UNP W8VZW3
A	-11	THR	-	expression tag	UNP W8VZW3
A	-10	GLY	-	expression tag	UNP W8VZW3
A	-9	GLY	-	expression tag	UNP W8VZW3
A	-8	GLN	-	expression tag	UNP W8VZW3
A	-7	GLN	-	expression tag	UNP W8VZW3
A	-6	MET	-	expression tag	UNP W8VZW3
A	-5	GLY	-	expression tag	UNP W8VZW3
A	-4	ARG	-	expression tag	UNP W8VZW3
A	-3	ASP	-	expression tag	UNP W8VZW3
A	-2	PRO	-	expression tag	UNP W8VZW3
A	-1	ASN	-	expression tag	UNP W8VZW3
A	0	SER	-	expression tag	UNP W8VZW3
A	273	LEU	-	expression tag	UNP W8VZW3
A	274	GLU	-	expression tag	UNP W8VZW3
A	275	HIS	-	expression tag	UNP W8VZW3
A	276	HIS	-	expression tag	UNP W8VZW3
A	277	HIS	-	expression tag	UNP W8VZW3
A	278	HIS	-	expression tag	UNP W8VZW3
A	279	HIS	-	expression tag	UNP W8VZW3
A	280	HIS	-	expression tag	UNP W8VZW3

- Molecule 2 is RETINAL (three-letter code: RET) (formula: C₂₀H₂₈O).

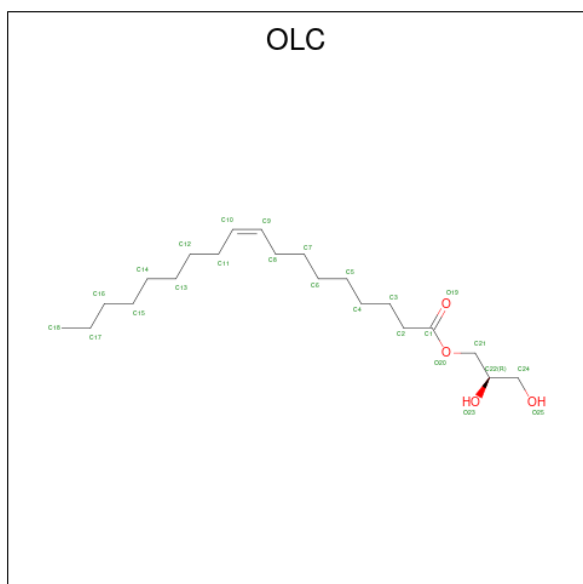


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C 20 20	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

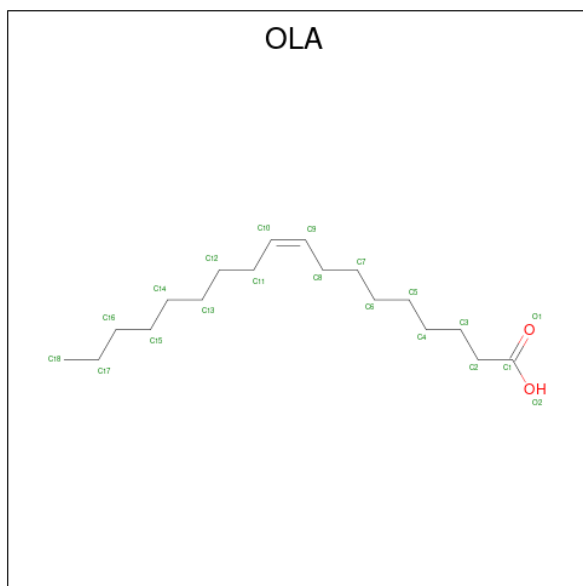
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0

- Molecule 4 is (2R)-2,3-dihydroxypropyl (9Z)-octadec-9-enoate (three-letter code: OLC) (formula: C₂₁H₄₀O₄).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 25 21 4	0	0
4	A	1	Total C O 18 14 4	0	0
4	A	1	Total C O 11 7 4	0	0

- Molecule 5 is OLEIC ACID (three-letter code: OLA) (formula: $C_{18}H_{34}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 20 18 2	0	0
5	A	1	Total C O 20 18 2	0	0
5	A	1	Total C O 13 11 2	0	0
5	A	1	Total C 14 14	0	0


- Molecule 6 is water.

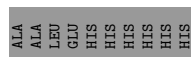
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	64	Total O 64 64	0	0

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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: Chloride pumping rhodopsin

Chain A:  83% 6% 11%



4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	104.48Å 51.18Å 78.33Å 90.00° 131.78° 90.00°	Depositor
Resolution (Å)	13.45 – 1.80 14.40 – 1.60	Depositor EDS
% Data completeness (in resolution range)	92.6 (13.45-1.80) 90.1 (14.40-1.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.69 (at 1.60Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.273 , 0.319 0.280 , 0.325	Depositor DCC
R_{free} test set	1866 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	11.6	Xtrriage
Anisotropy	0.922	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.034 for h,-k,-h-l	Xtrriage
F_o, F_c correlation	0.85	EDS
Total number of atoms	2290	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.17% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: OLA, CL, RET, OLC

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	0.30	0/2149	0.44	0/2931

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	2084	0	2127	11	0
2	A	20	0	27	4	0
3	A	1	0	0	2	0
4	A	54	0	74	0	0
5	A	67	0	107	1	0
6	A	64	0	0	0	0
All	All	2290	0	2335	15	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 3.

All (15) 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:49:ILE:HG23	1:A:112:ILE:HD13	1.67	0.74
1:A:98:ASN:ND2	3:A:302:CL:CL	2.64	0.64
2:A:301:RET:H8	2:A:301:RET:H161	1.81	0.62
2:A:301:RET:H8	2:A:301:RET:H171	1.87	0.56
1:A:58:MET:HE1	1:A:236:VAL:HG22	1.91	0.53
1:A:103:ILE:HD12	1:A:136:ILE:HD11	1.92	0.52
1:A:49:ILE:CG2	1:A:112:ILE:HD13	2.40	0.49
1:A:12:ALA:O	1:A:16:GLU:HG3	2.13	0.48
2:A:301:RET:H171	2:A:301:RET:C8	2.45	0.46
2:A:301:RET:H161	2:A:301:RET:C8	2.45	0.43
1:A:64:ILE:HG13	1:A:98:ASN:CG	2.41	0.41
1:A:23[A]:THR:HG22	5:A:306:OLA:H31	2.03	0.41
1:A:43:VAL:HG21	1:A:51[A]:THR:HG21	2.03	0.40
1:A:102:THR:HG21	3:A:302:CL:CL	2.58	0.40
1:A:193:VAL:O	1:A:197:MET:HG3	2.21	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	265/296 (90%)	261 (98%)	3 (1%)	1 (0%)	34 21

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	115	ASN

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	221/242 (91%)	221 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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	3	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](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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	OLC	A	303	-	24,24,24	0.80	2 (8%)	25,25,25	0.91	1 (4%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	RET	A	301	1	20,20,21	0.69	0	27,27,28	1.69	5 (18%)
4	OLC	A	307	-	10,10,24	1.19	2 (20%)	11,11,25	1.10	1 (9%)
5	OLA	A	305	-	16,19,19	0.70	1 (6%)	15,19,19	0.66	0
5	OLA	A	309	-	13,13,19	0.80	1 (7%)	12,12,19	0.49	0
4	OLC	A	304	-	17,17,24	0.96	2 (11%)	18,18,25	1.04	1 (5%)
5	OLA	A	306	-	16,19,19	0.73	1 (6%)	15,19,19	0.53	0
5	OLA	A	308	-	9,12,19	0.84	1 (11%)	8,12,19	0.77	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	OLC	A	303	-	-	12/24/24/24	-
2	RET	A	301	1	-	0/13/30/31	0/1/1/1
4	OLC	A	307	-	-	7/10/10/24	-
5	OLA	A	305	-	-	8/15/17/17	-
5	OLA	A	309	-	-	5/11/11/17	-
4	OLC	A	304	-	-	8/17/17/24	-
5	OLA	A	306	-	-	10/15/17/17	-
5	OLA	A	308	-	-	5/8/10/17	-

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	306	OLA	C10-C9	2.75	1.47	1.31
5	A	309	OLA	C10-C9	2.72	1.47	1.31
5	A	305	OLA	C10-C9	2.70	1.47	1.31
4	A	304	OLC	O20-C1	2.49	1.40	1.33
5	A	308	OLA	C9-C10	2.41	1.47	1.29
4	A	303	OLC	O20-C1	2.39	1.40	1.33
4	A	307	OLC	O20-C1	2.36	1.40	1.33
4	A	307	OLC	O20-C21	-2.15	1.40	1.45
4	A	304	OLC	O20-C21	-2.07	1.40	1.45
4	A	303	OLC	O20-C21	-2.03	1.40	1.45

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	301	RET	C11-C12-C13	-4.25	114.47	126.42
2	A	301	RET	C18-C5-C6	-3.28	120.85	124.53
2	A	301	RET	C11-C10-C9	-2.74	123.40	127.31
4	A	304	OLC	O20-C1-C2	2.62	120.14	111.91
2	A	301	RET	C3-C4-C5	-2.59	109.46	114.08
4	A	307	OLC	O20-C1-C2	2.57	119.97	111.91
2	A	301	RET	C7-C8-C9	-2.50	122.46	126.23
4	A	303	OLC	O20-C1-C2	2.50	119.74	111.91

There are no chirality outliers.

All (55) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	306	OLA	C1-C2-C3-C4
4	A	304	OLC	O19-C1-O20-C21
4	A	304	OLC	C2-C1-O20-C21
4	A	303	OLC	C2-C1-O20-C21
4	A	307	OLC	C2-C1-O20-C21
4	A	304	OLC	O20-C21-C22-C24
4	A	304	OLC	C1-C2-C3-C4
4	A	303	OLC	C1-C2-C3-C4
4	A	303	OLC	O19-C1-O20-C21
4	A	307	OLC	O19-C1-O20-C21
4	A	304	OLC	O20-C21-C22-O23
5	A	305	OLA	C11-C10-C9-C8
4	A	303	OLC	C11-C12-C13-C14
5	A	306	OLA	C14-C15-C16-C17
4	A	304	OLC	C3-C4-C5-C6
4	A	303	OLC	C6-C7-C8-C9
5	A	306	OLA	C3-C4-C5-C6
5	A	308	OLA	C5-C6-C7-C8
4	A	307	OLC	C1-C2-C3-C4
5	A	309	OLA	C6-C7-C8-C9
5	A	306	OLA	C5-C6-C7-C8
5	A	305	OLA	C10-C11-C12-C13
5	A	306	OLA	C6-C7-C8-C9
4	A	303	OLC	O20-C21-C22-C24
5	A	309	OLA	C15-C16-C17-C18
4	A	304	OLC	C5-C6-C7-C8
5	A	309	OLA	C5-C6-C7-C8
4	A	303	OLC	C15-C16-C17-C18
5	A	306	OLA	C11-C10-C9-C8
4	A	303	OLC	O20-C21-C22-O23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	A	308	OLA	C11-C10-C9-C8
5	A	306	OLA	C11-C12-C13-C14
5	A	305	OLA	C5-C6-C7-C8
5	A	306	OLA	C4-C5-C6-C7
5	A	309	OLA	C14-C15-C16-C17
5	A	305	OLA	C1-C2-C3-C4
5	A	306	OLA	C12-C13-C14-C15
5	A	305	OLA	C4-C5-C6-C7
4	A	307	OLC	O20-C21-C22-O23
4	A	303	OLC	C5-C6-C7-C8
4	A	303	OLC	C21-C22-C24-O25
5	A	308	OLA	C6-C7-C8-C9
5	A	305	OLA	C11-C12-C13-C14
5	A	309	OLA	C11-C10-C9-C8
4	A	304	OLC	C7-C8-C9-C10
5	A	306	OLA	C9-C10-C11-C12
4	A	303	OLC	C9-C10-C11-C12
4	A	303	OLC	C7-C8-C9-C10
5	A	308	OLA	C7-C8-C9-C10
5	A	305	OLA	C9-C10-C11-C12
5	A	305	OLA	C2-C3-C4-C5
4	A	307	OLC	C21-C22-C24-O25
5	A	308	OLA	C4-C5-C6-C7
4	A	307	OLC	O20-C1-C2-C3
4	A	307	OLC	O19-C1-C2-C3

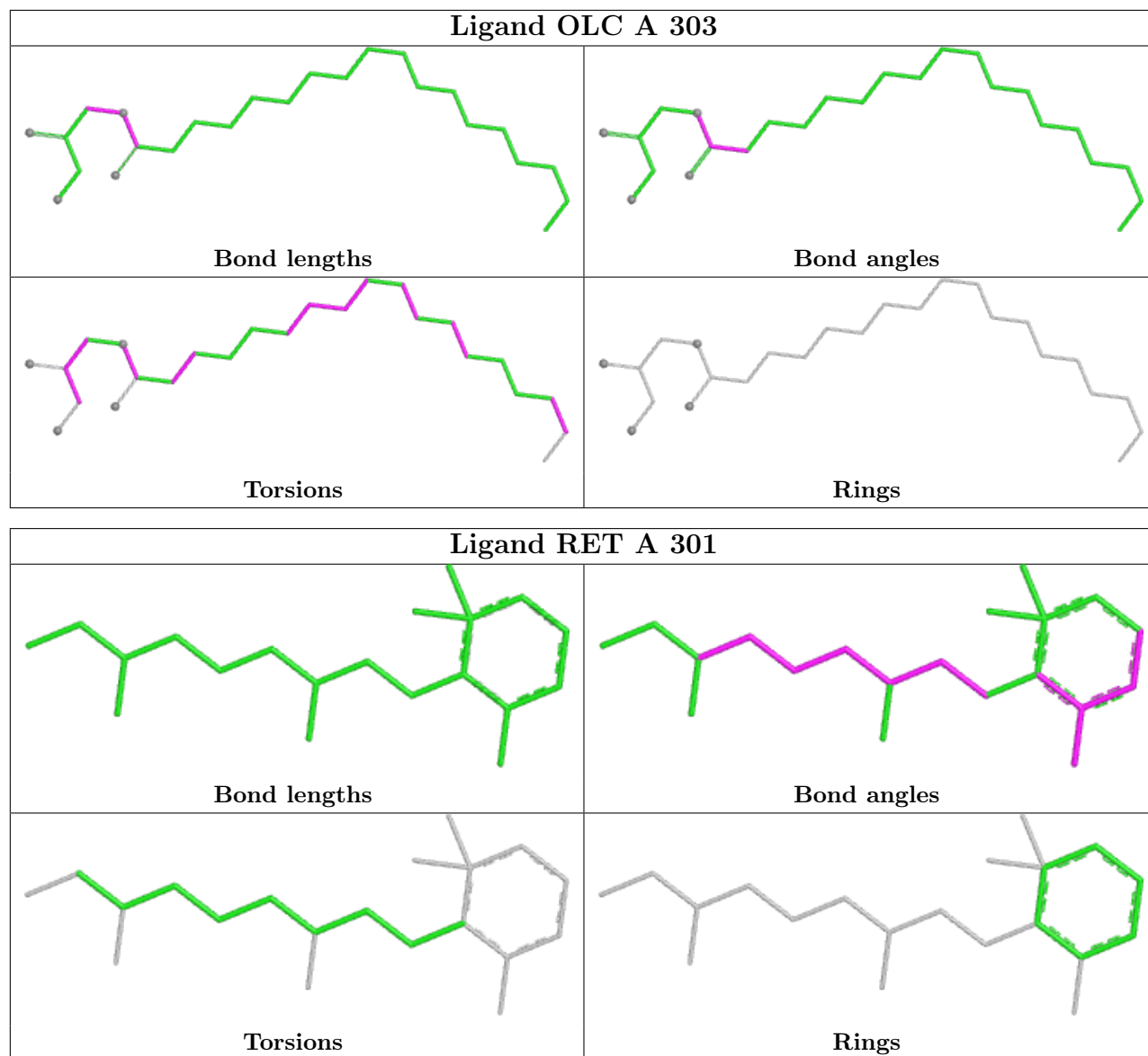
There are no ring outliers.

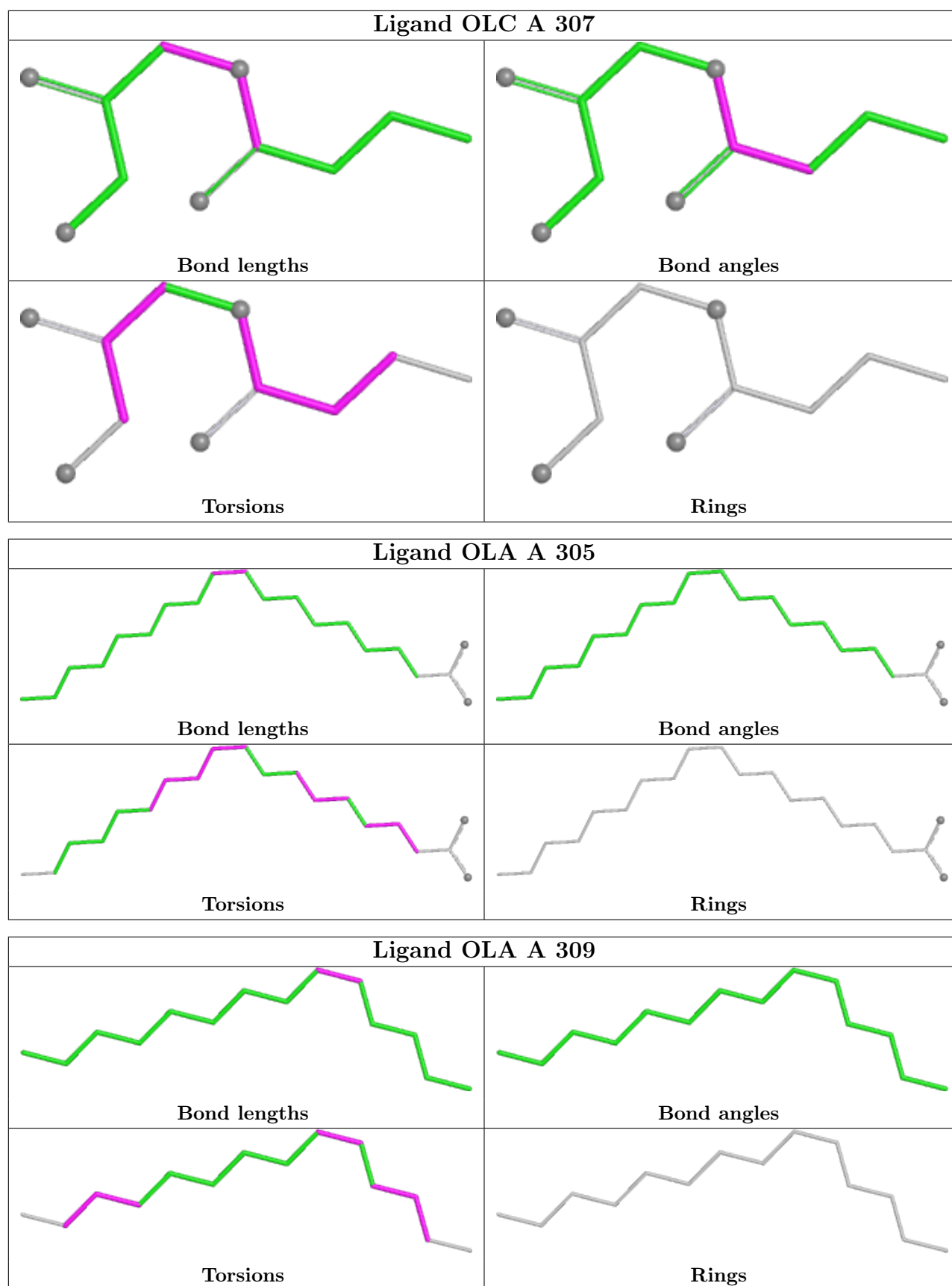
2 monomers are involved in 5 short contacts:

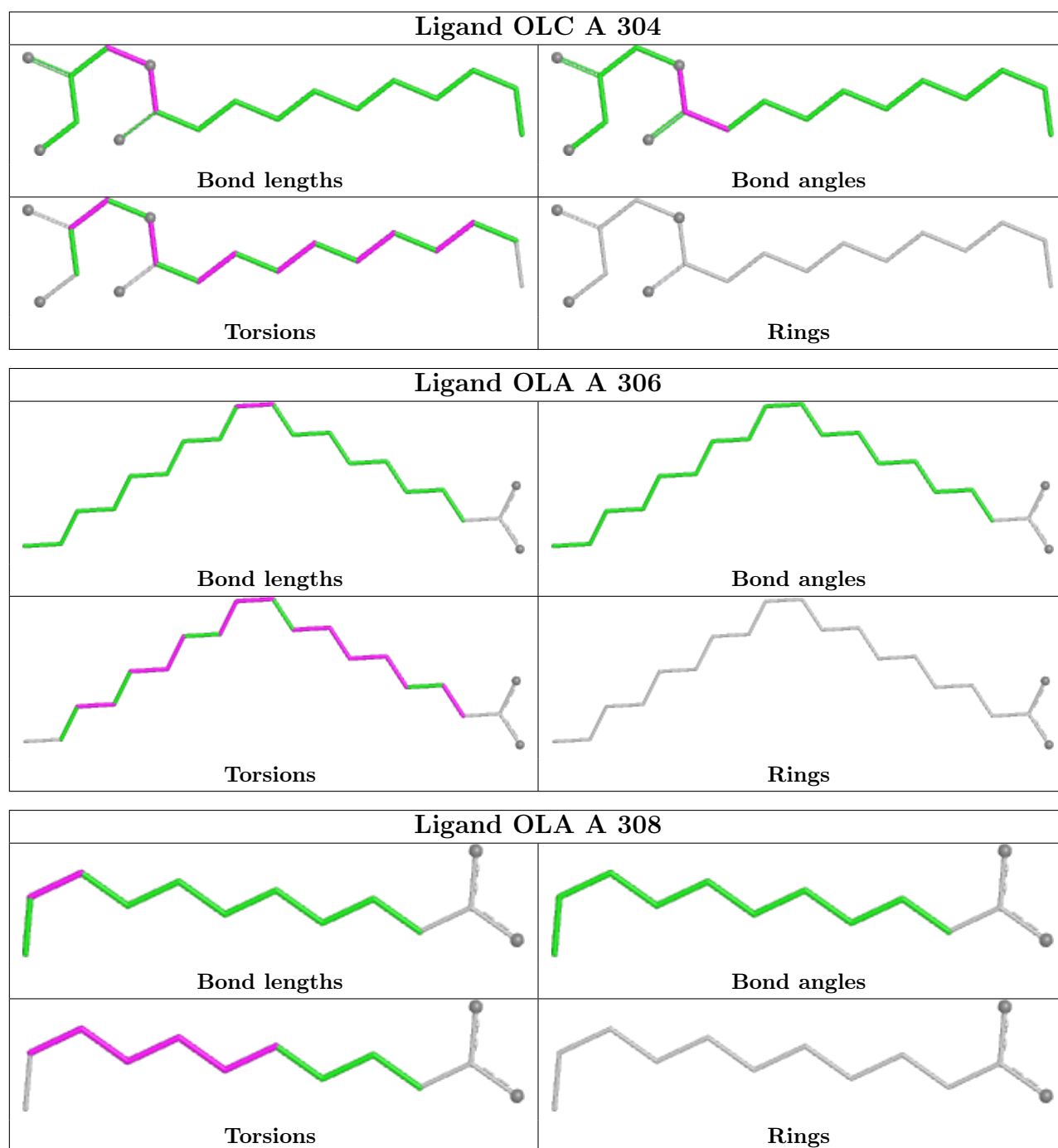
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	301	RET	4	0
5	A	306	OLA	1	0

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](#)

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.

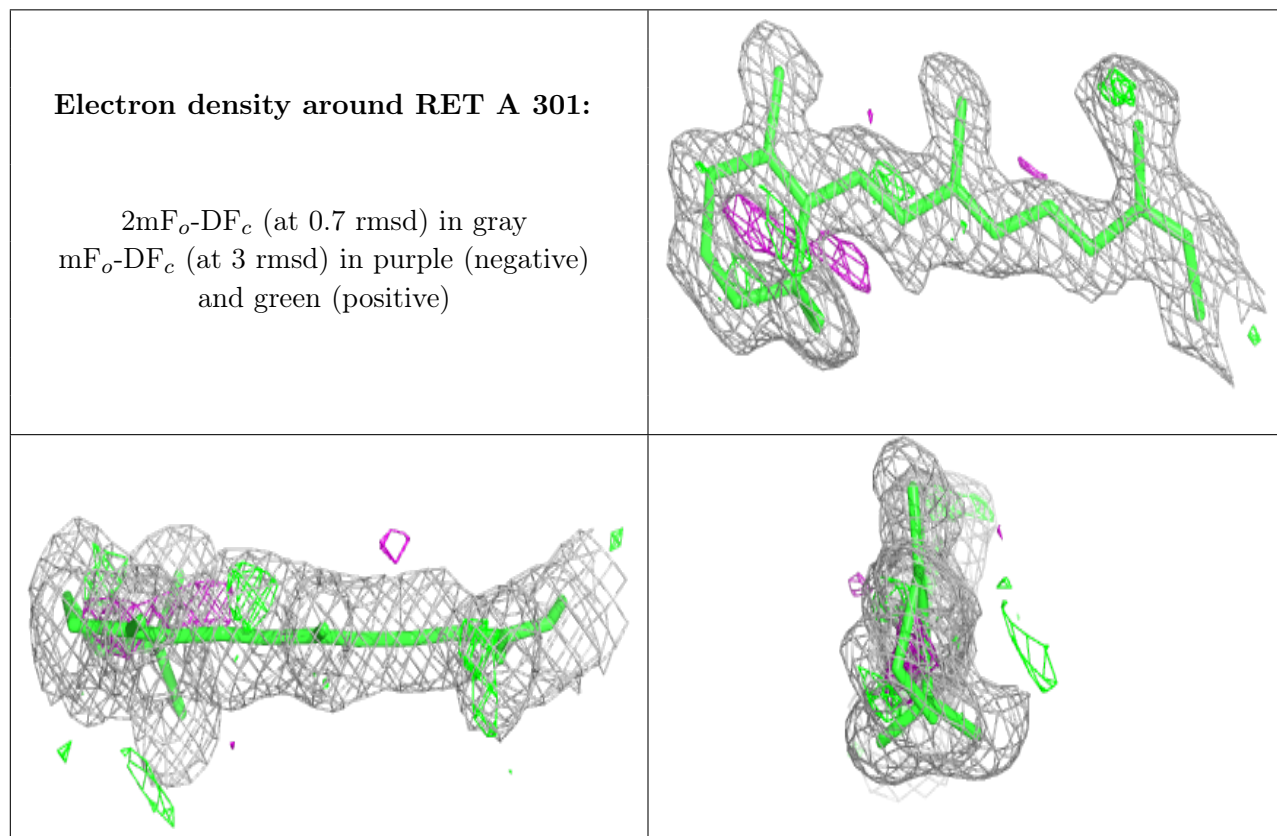
6.3 Carbohydrates [i](#)

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

6.4 Ligands [i](#)

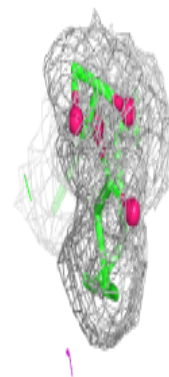
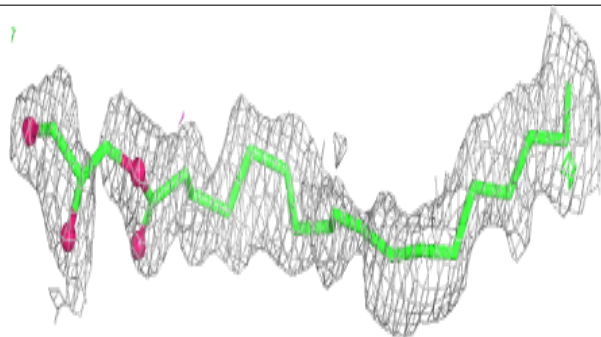
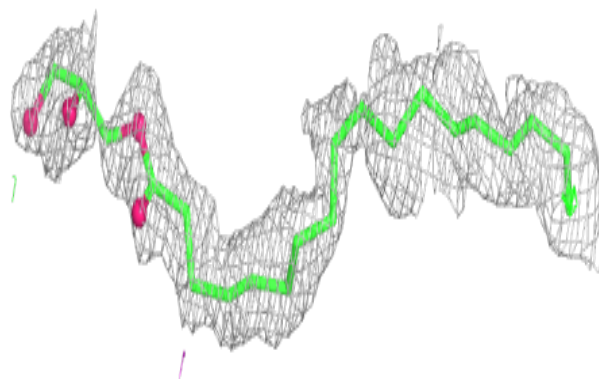
Unable to reproduce the depositors R factor - this section is therefore empty.

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.

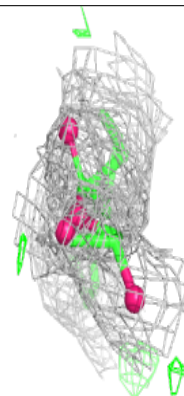
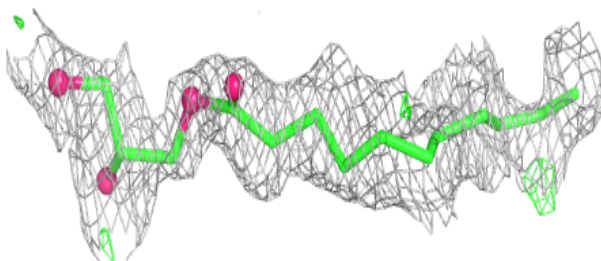
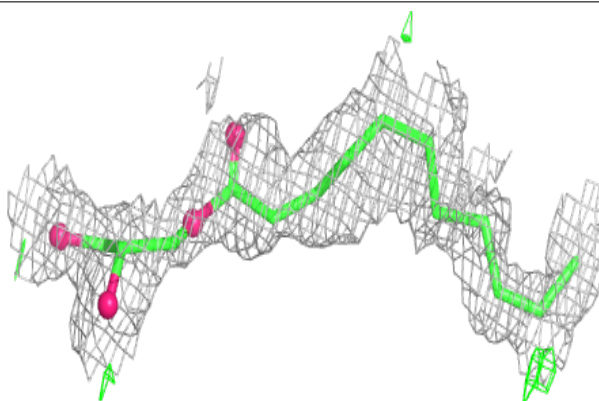


Electron density around OLC A 303:

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

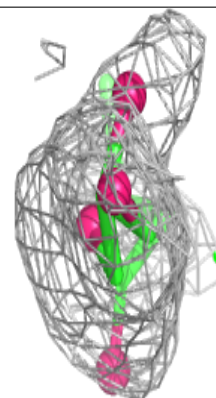
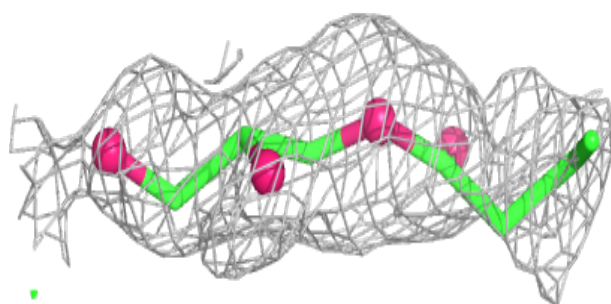
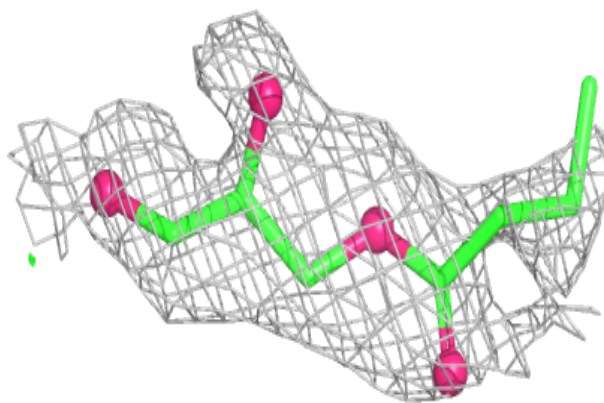
**Electron density around OLC A 304:**

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

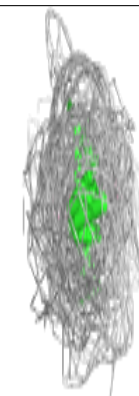
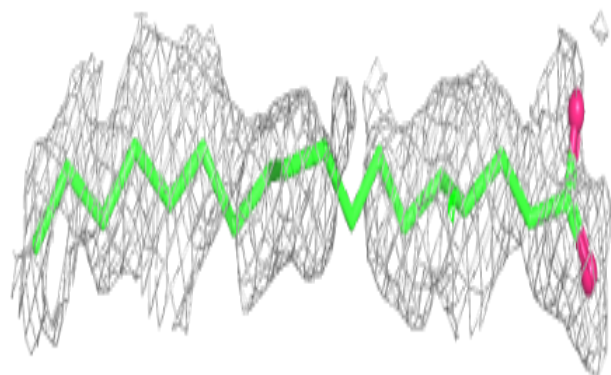
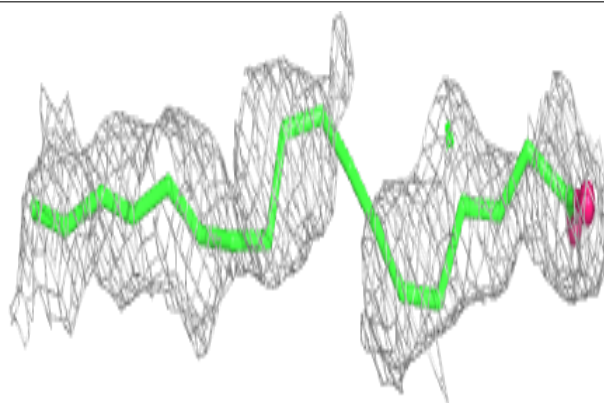


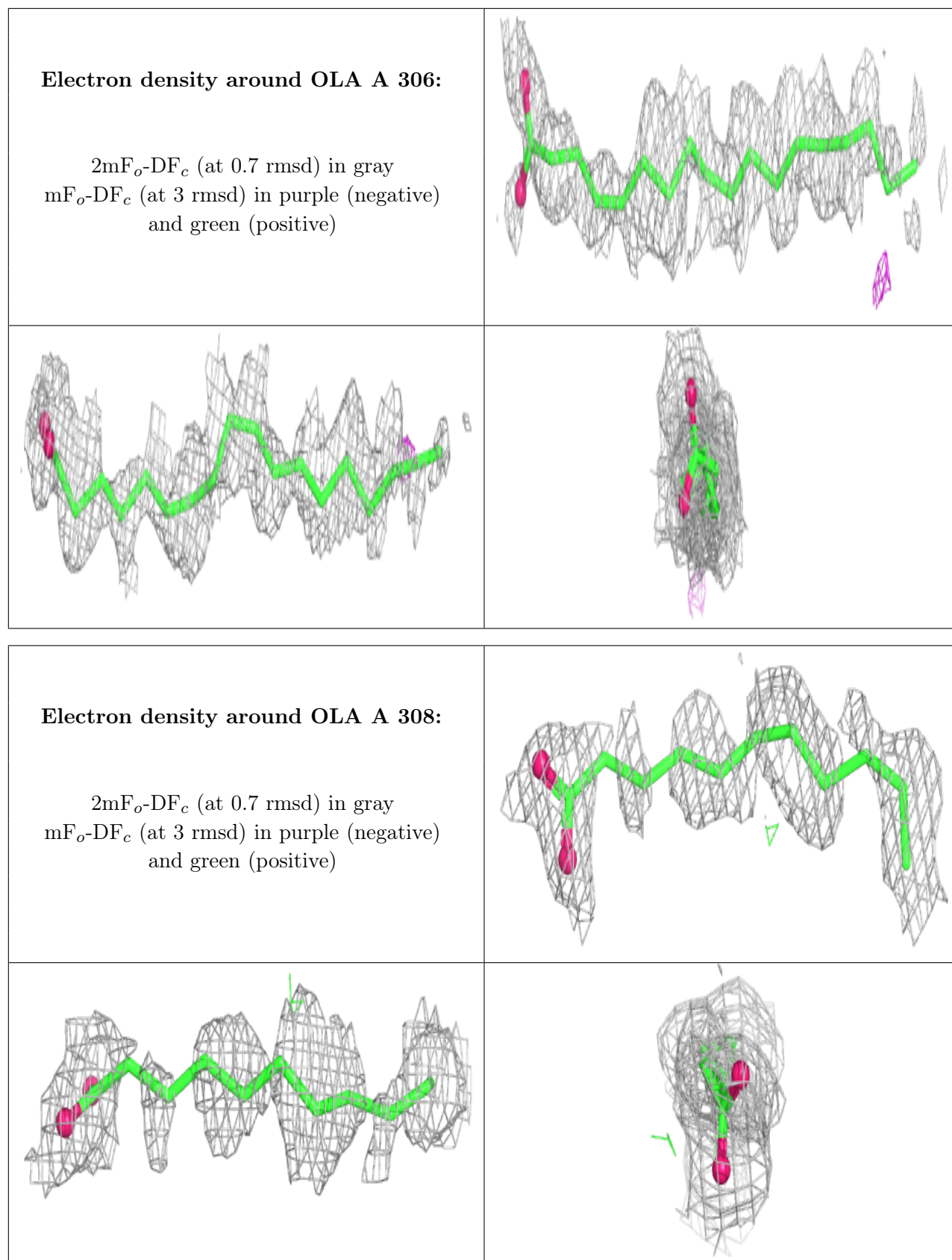
Electron density around OLC A 307:

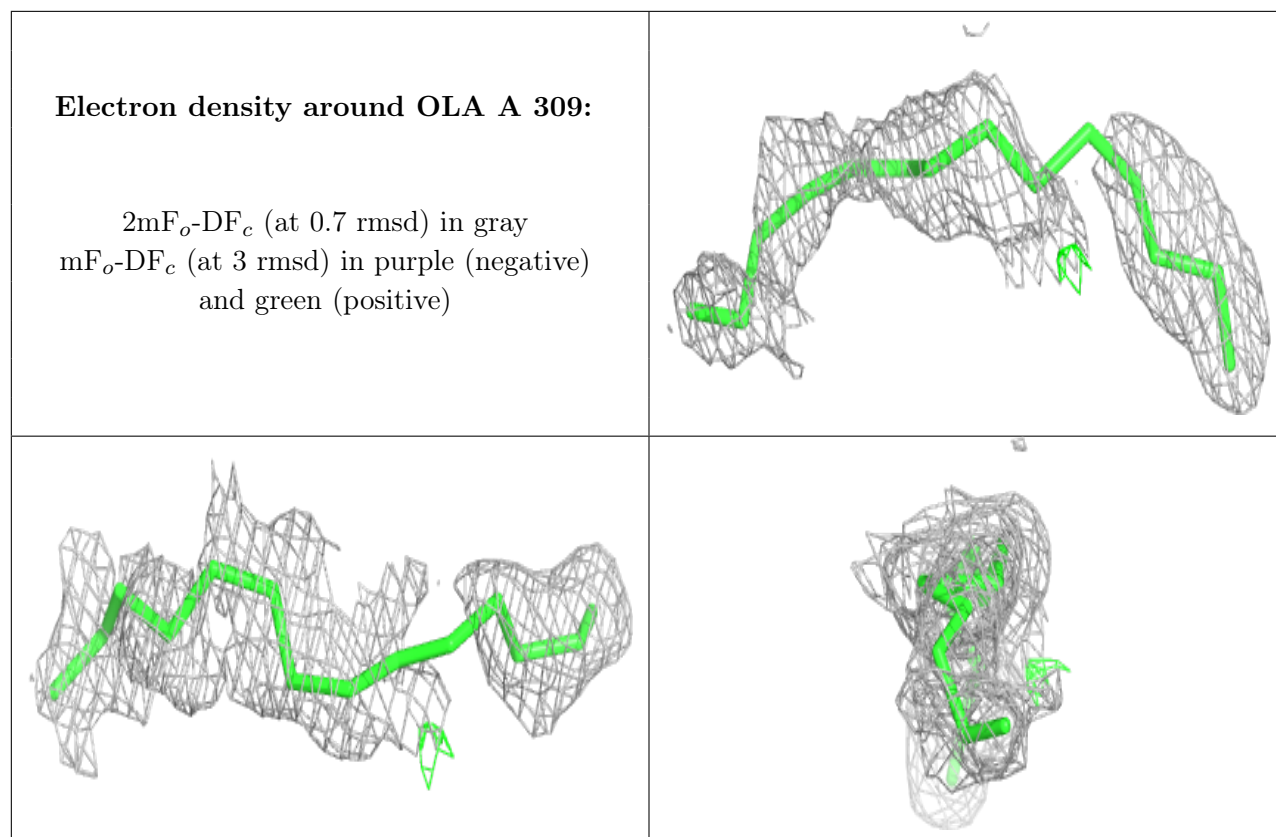
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

**Electron density around OLA A 305:**

$2mF_o-DF_c$ (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.