



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 6, 2024 – 11:00 AM EDT

PDB ID : 1KDG  
Title : Crystal structure of the flavin domain of cellobiose dehydrogenase  
Authors : Hallberg, B.M.; Henriksson, G.; Pettersson, G.; Divne, C.  
Deposited on : 2001-11-13  
Resolution : 1.50 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)  
Xtrriage (Phenix) : 1.20.1  
EDS : **FAILED**  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

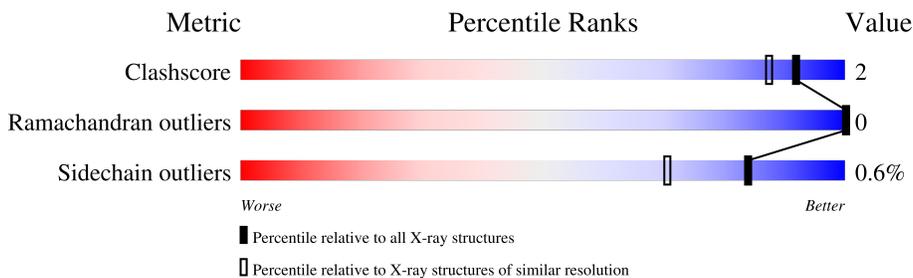
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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))
Clashscore	180529	4048 (1.50-1.50)
Ramachandran outliers	177936	3970 (1.50-1.50)
Sidechain outliers	177891	3967 (1.50-1.50)

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.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	546	95%
1	B	546	95%

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
6	EMT	A	6907	-	X	-	-

## 2 Entry composition [i](#)

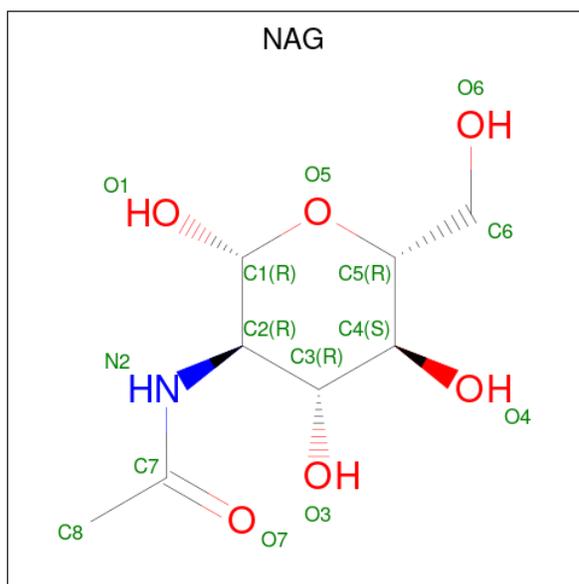
There are 8 unique types of molecules in this entry. The entry contains 9548 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 cellobiose dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	541	Total	C	N	O	S	0	4	0
			4075	2585	693	789	8			
1	B	546	Total	C	N	O	S	0	7	0
			4110	2605	698	799	8			

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



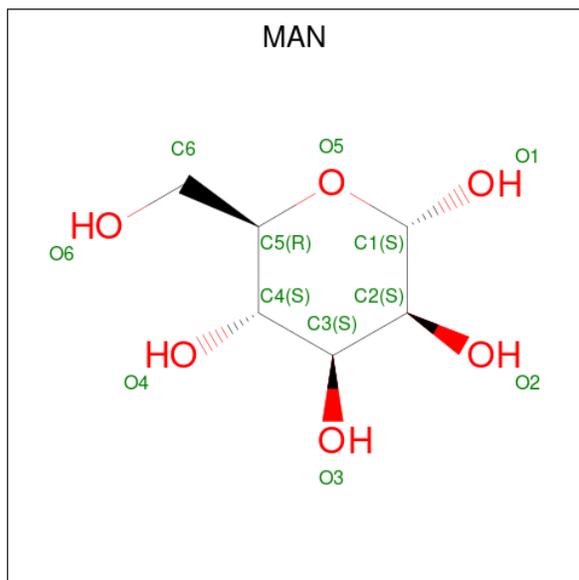
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

Continued from previous page...

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

- Molecule 3 is alpha-D-mannopyranose (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).

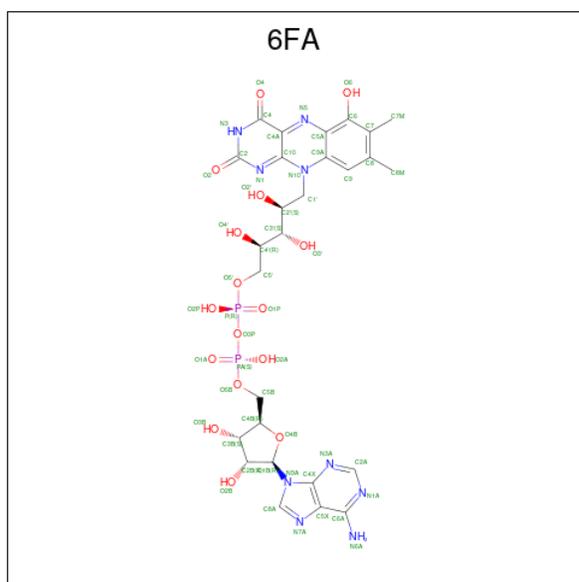


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	11	6	5	0	0
3	B	1	11	6	5	0	0
3	B	1	11	6	5	0	0
3	B	1	11	6	5	0	0

- Molecule 4 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

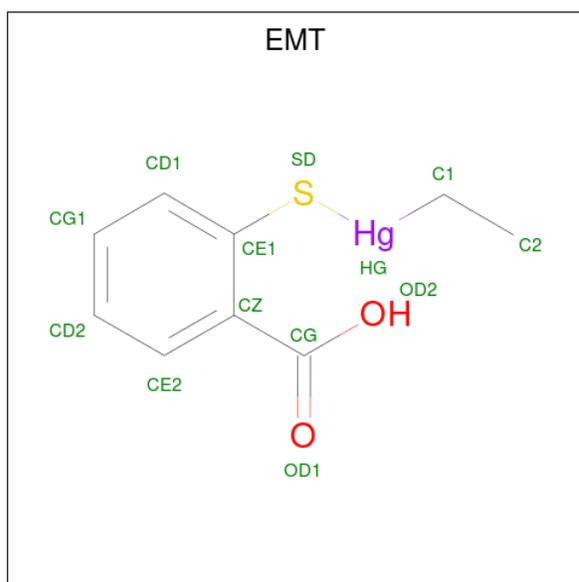
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Hg		
4	A	6	6	6	0	0
4	B	3	3	3	0	0

- Molecule 5 is 6-HYDROXY-FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: 6FA) (formula: C<sub>27</sub>H<sub>33</sub>N<sub>9</sub>O<sub>16</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	O			P
5	A	1	54	27	9	16	2	0	0
5	B	1	54	27	9	16	2	0	0

- Molecule 6 is 2-(ETHYLMERCURI-THIO)-BENZOIC ACID (three-letter code: EMT) (formula: C<sub>9</sub>H<sub>10</sub>HgO<sub>2</sub>S).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	Hg	O			S
6	A	1	11	7	1	2	1	0	0

- Molecule 7 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	1	Total X 1 1	0	0

- Molecule 8 is water.

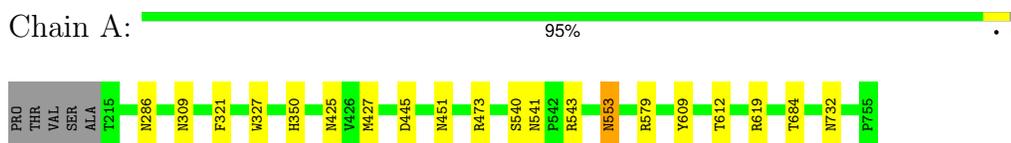
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	587	Total O 587 587	0	0
8	B	533	Total O 533 533	0	0

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

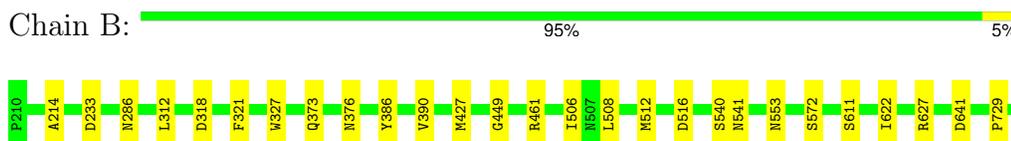
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

Note EDS failed to run properly.

- Molecule 1: cellobiose dehydrogenase



- Molecule 1: cellobiose dehydrogenase



## 4 Data and refinement statistics [i](#)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	185.07Å 185.07Å 81.03Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	23.00 – 1.50	Depositor
% Data completeness (in resolution range)	99.6 (23.00-1.50)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.63 (at 1.50Å)	Xtrriage
Refinement program	REFMAC 5.0	Depositor
R, $R_{free}$	0.133 , 0.169	Depositor
Wilson B-factor (Å <sup>2</sup> )	12.4	Xtrriage
Anisotropy	0.058	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.022 for h,-h-k,-l	Xtrriage
Total number of atoms	9548	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.20% 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: 6FA, EMT, UNX, HG, MAN, 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	0.98	1/4207 (0.0%)	0.89	4/5767 (0.1%)
1	B	1.01	3/4258 (0.1%)	0.93	8/5837 (0.1%)
All	All	0.99	4/8465 (0.0%)	0.91	12/11604 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	427	MET	CG-SD	-9.19	1.57	1.81
1	A	427	MET	CG-SD	-8.49	1.59	1.81
1	B	512	MET	CG-SD	-5.93	1.65	1.81
1	B	746	ALA	CA-CB	-5.02	1.41	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	461	ARG	NE-CZ-NH1	7.34	123.97	120.30
1	B	461	ARG	NE-CZ-NH2	-6.74	116.93	120.30
1	B	318	ASP	CB-CG-OD2	6.59	124.23	118.30
1	B	516	ASP	CB-CG-OD2	6.18	123.86	118.30
1	B	572[A]	SER	CA-CB-OG	-6.14	94.61	111.20

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	4075	0	3950	13	0
1	B	4110	0	3986	10	0
2	A	42	0	38	0	0
2	B	28	0	26	1	0
3	A	11	0	10	0	0
3	B	33	0	30	0	0
4	A	6	0	0	0	0
4	B	3	0	0	0	0
5	A	54	0	28	2	0
5	B	54	0	28	0	0
6	A	11	0	4	0	0
7	B	1	0	0	0	0
8	A	587	0	0	5	0
8	B	533	0	0	3	0
All	All	9548	0	8100	24	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 24 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:B:286:ASN:ND2	8:B:3417:HOH:O	2.03	0.91
1:A:286:ASN:ND2	8:A:1417:HOH:O	1.86	0.86
1:A:540:SER:OG	1:A:541[B]:ASN:ND2	2.22	0.73
1:A:541[B]:ASN:OD1	8:A:2092:HOH:O	2.16	0.58
1:B:540:SER:OG	1:B:541[B]:ASN:ND2	2.39	0.54

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	543/546 (100%)	522 (96%)	21 (4%)	0	100	100
1	B	551/546 (101%)	534 (97%)	17 (3%)	0	100	100
All	All	1094/1092 (100%)	1056 (96%)	38 (4%)	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	443/443 (100%)	440 (99%)	3 (1%)	81	66
1	B	450/443 (102%)	448 (100%)	2 (0%)	89	79
All	All	893/886 (101%)	888 (99%)	5 (1%)	84	70

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	451	ASN
1	A	553	ASN
1	A	609	TYR
1	B	376	ASN
1	B	553	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	350	HIS
1	B	521	ASN
1	B	442	GLN
1	B	553	ASN
1	A	451	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 22 ligands modelled in this entry, 9 are monoatomic and 1 is unknown - leaving 12 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
2	NAG	B	803	1	14,14,15	1.07	1 (7%)	17,19,21	1.40	3 (17%)
3	MAN	A	805	1	11,11,12	0.53	0	15,15,17	1.11	1 (6%)
2	NAG	A	803	1	14,14,15	0.92	1 (7%)	17,19,21	1.15	1 (5%)
3	MAN	B	805	1	11,11,12	0.67	0	15,15,17	1.01	1 (6%)
2	NAG	A	804	1	14,14,15	0.80	1 (7%)	17,19,21	1.50	3 (17%)
5	6FA	A	6801	-	55,59,59	1.87	13 (23%)	68,91,91	1.59	11 (16%)
6	EMT	A	6907	4	10,11,13	3.53	7 (70%)	13,14,16	4.00	6 (46%)
5	6FA	B	7801	-	55,59,59	2.21	21 (38%)	68,91,91	1.82	17 (25%)
2	NAG	B	804	1	14,14,15	0.75	0	17,19,21	1.80	4 (23%)
3	MAN	B	806	1	11,11,12	0.48	0	15,15,17	1.02	0
2	NAG	A	802	1	14,14,15	1.29	3 (21%)	17,19,21	1.47	4 (23%)
3	MAN	B	807	1	11,11,12	0.62	0	15,15,17	1.07	1 (6%)

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	NAG	B	803	1	-	1/6/23/26	0/1/1/1
3	MAN	A	805	1	-	2/2/19/22	0/1/1/1
2	NAG	A	803	1	-	0/6/23/26	0/1/1/1
3	MAN	B	805	1	-	0/2/19/22	0/1/1/1
2	NAG	A	804	1	-	3/6/23/26	0/1/1/1
5	6FA	A	6801	-	-	3/30/50/50	0/6/6/6
6	EMT	A	6907	4	-	0/4/6/8	0/1/1/1
5	6FA	B	7801	-	-	4/30/50/50	0/6/6/6
2	NAG	B	804	1	-	2/6/23/26	0/1/1/1
3	MAN	B	806	1	-	2/2/19/22	0/1/1/1
2	NAG	A	802	1	-	0/6/23/26	0/1/1/1
3	MAN	B	807	1	-	2/2/19/22	0/1/1/1

The worst 5 of 47 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	6907	EMT	OD2-CG	-8.10	1.06	1.30
5	A	6801	6FA	C9A-C5A	-5.54	1.33	1.41
5	B	7801	6FA	C4A-N5	4.78	1.41	1.30
5	B	7801	6FA	C2B-C3B	-4.64	1.40	1.53
6	A	6907	EMT	OD1-CG	-4.61	1.09	1.22

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	6907	EMT	OD2-CG-OD1	-11.73	98.13	123.35
6	A	6907	EMT	OD2-CG-CZ	5.67	131.41	115.28
2	B	804	NAG	O5-C1-C2	-5.27	103.14	111.29
5	B	7801	6FA	C4A-C10-N10	5.04	123.69	116.48
5	A	6801	6FA	C4B-O4B-C1B	-4.52	105.78	109.92

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	804	NAG	C8-C7-N2-C2
2	A	804	NAG	O7-C7-N2-C2

*Continued on next page...*

*Continued from previous page...*

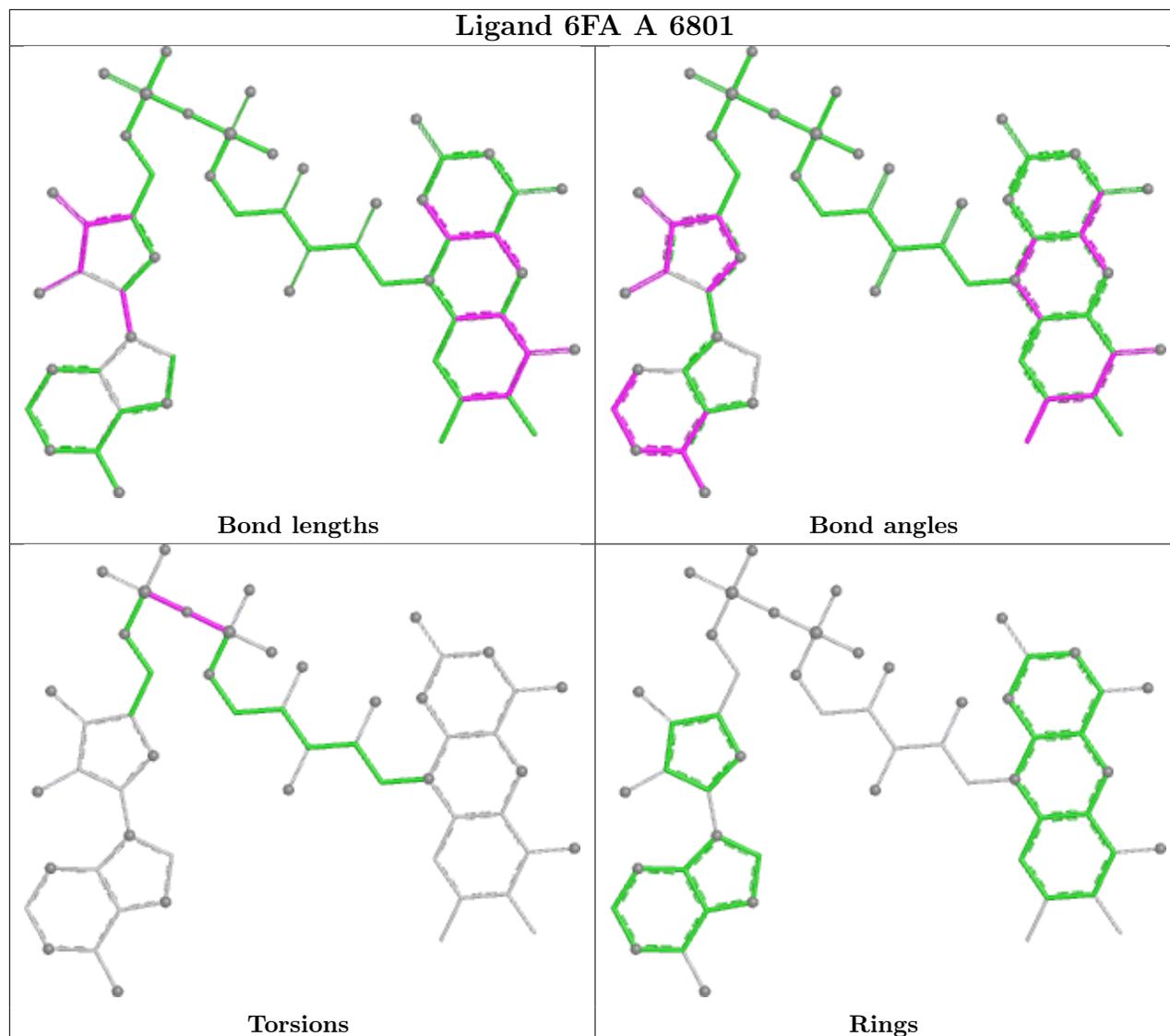
Mol	Chain	Res	Type	Atoms
3	B	807	MAN	O5-C5-C6-O6
3	A	805	MAN	O5-C5-C6-O6
3	B	806	MAN	O5-C5-C6-O6

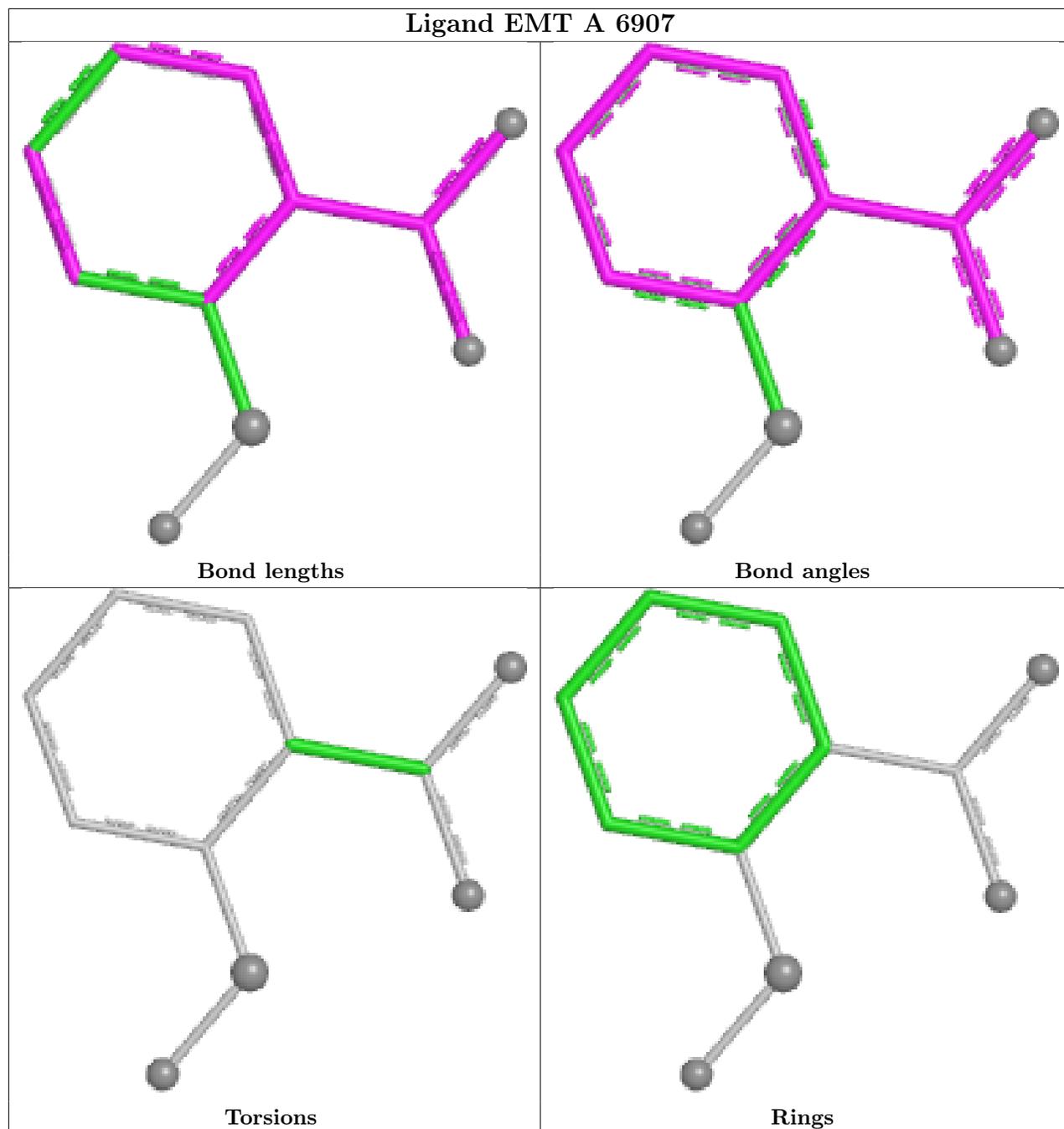
There are no ring outliers.

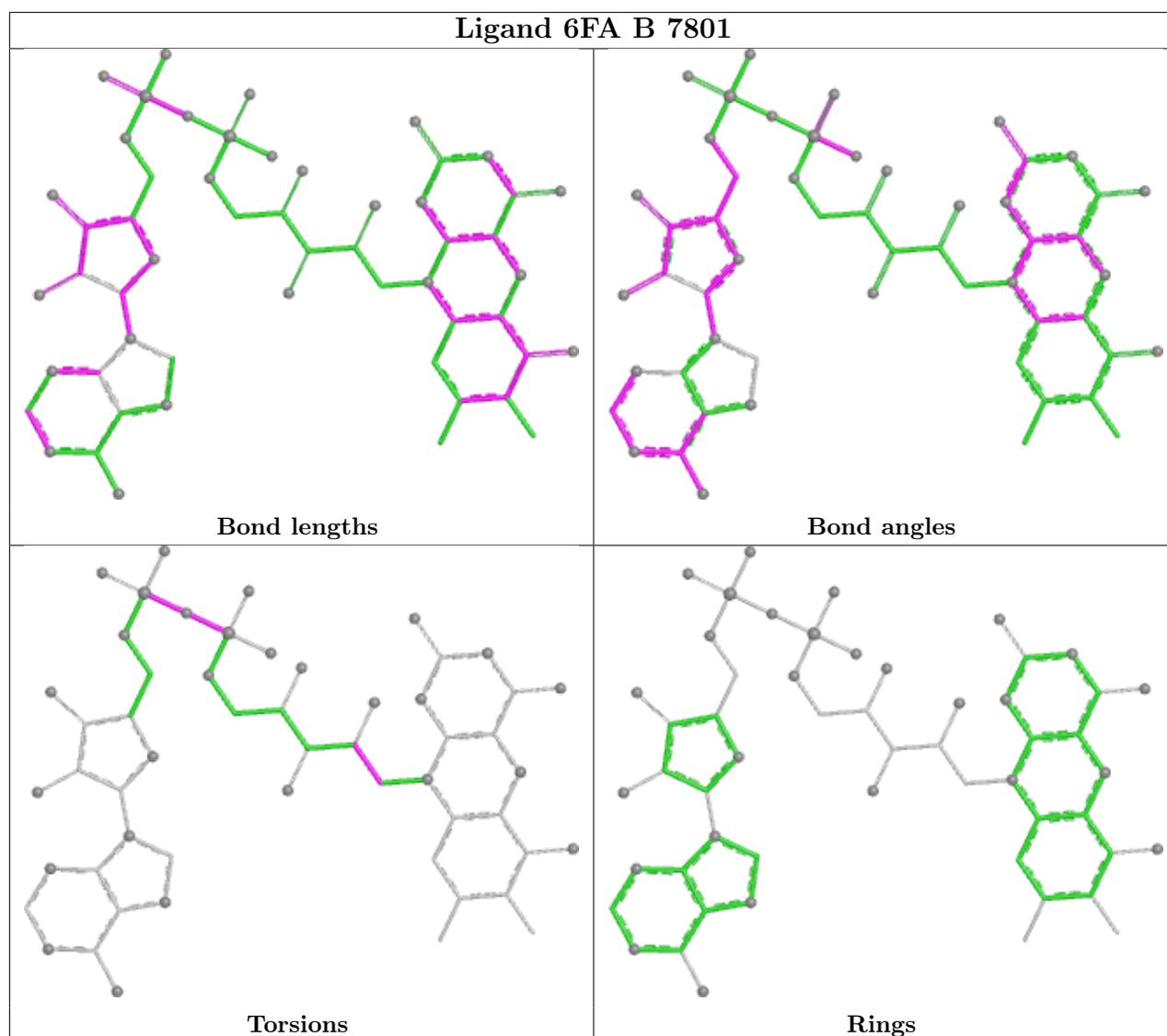
2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	6801	6FA	2	0
2	B	804	NAG	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](#)

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

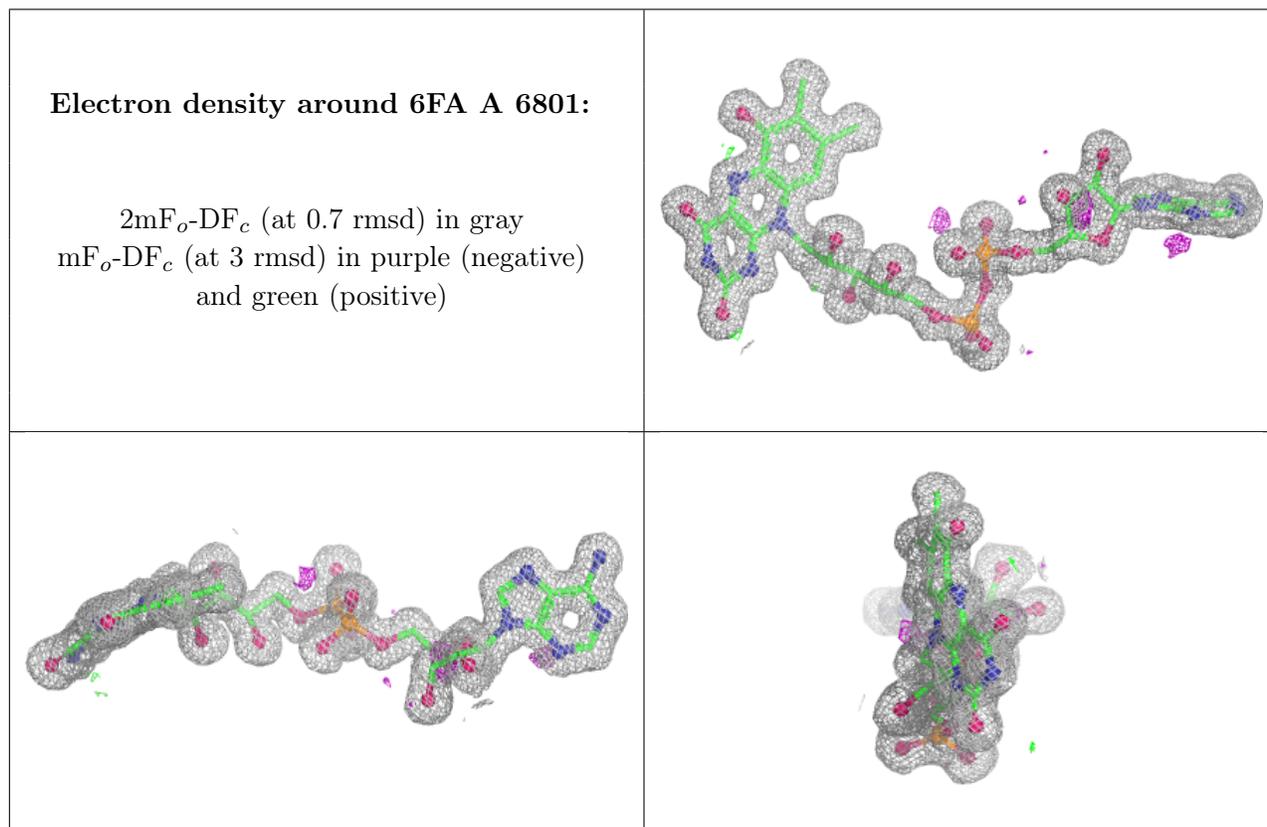
### 6.3 Carbohydrates [i](#)

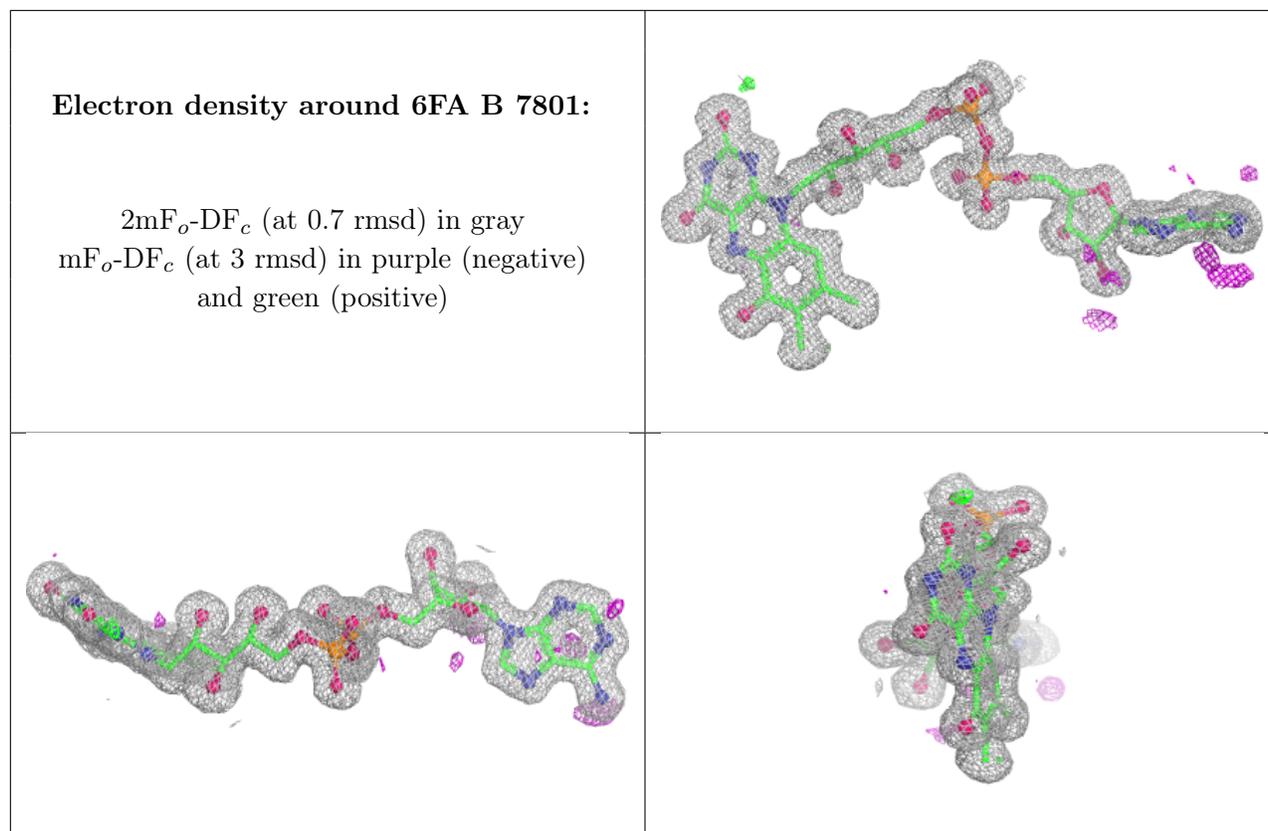
EDS failed to run properly - this section is therefore empty.

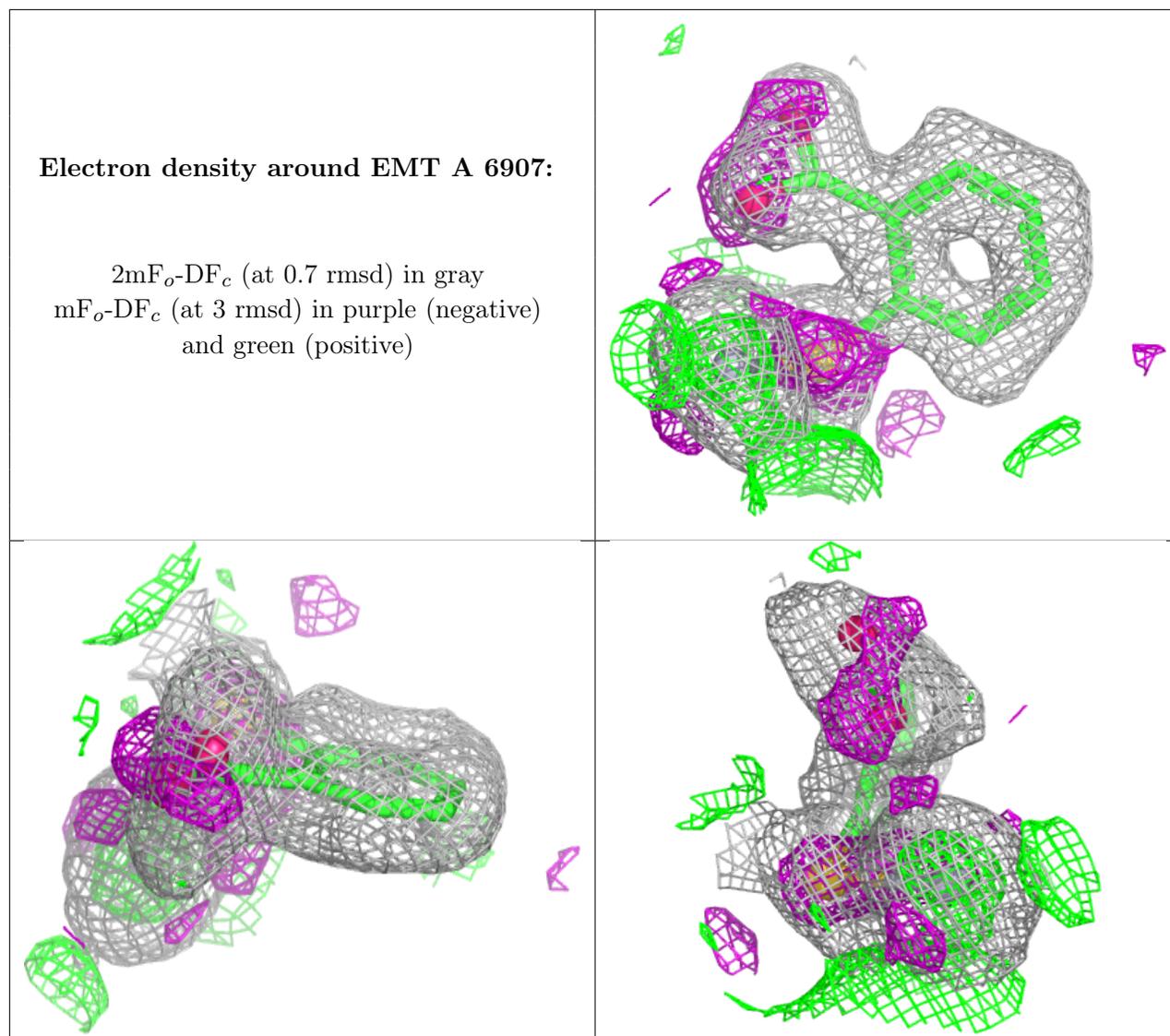
### 6.4 Ligands [i](#)

EDS failed to run properly - 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.







## 6.5 Other polymers [i](#)

EDS failed to run properly - this section is therefore empty.