



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 1, 2021 – 12:35 PM EDT

PDB ID : 5WEH  
Title : Cytochrome c oxidase from Rhodobacter sphaeroides in the reduced state  
Authors : Liu, J.; Ferguson-Miller, F.; Ling, Q.; Hiser, C.  
Deposited on : 2017-07-10  
Resolution : 3.45 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 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 : 2.22  
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) : 2.22

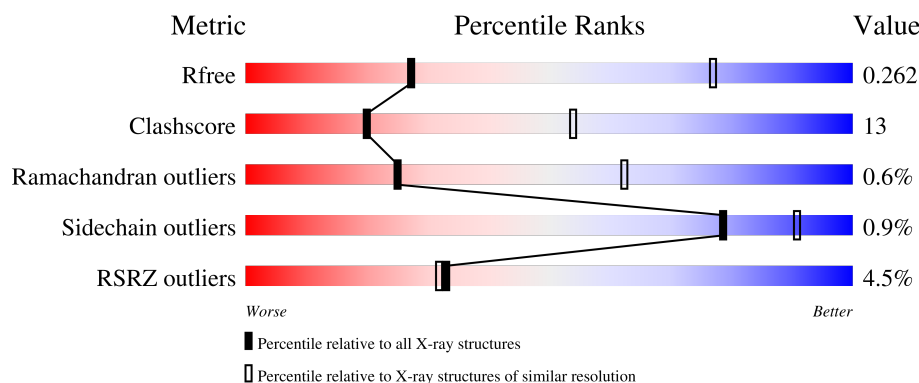
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*



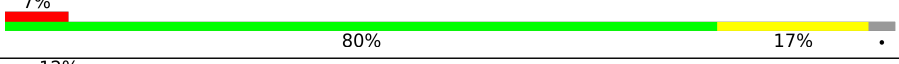


The reported resolution of this entry is 3.45 Å.

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	1291 (3.52-3.40)
Clashscore	141614	1372 (3.52-3.40)
Ramachandran outliers	138981	1337 (3.52-3.40)
Sidechain outliers	138945	1338 (3.52-3.40)
RSRZ outliers	127900	1205 (3.52-3.40)

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	566	
1	G	566	
2	B	262	
2	H	262	
3	C	266	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	I	266	
4	D	50	
4	J	50	

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
10	LMU	G	609	-	-	-	X
5	HEA	A	602	X	-	X	-
5	HEA	G	602	-	-	X	-
5	HEA	G	603	X	-	-	-
7	MG	G	605	-	-	-	X
9	3PE	C	301	-	-	-	X
9	3PE	I	303	-	-	-	X

## 2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 17909 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 Cytochrome c oxidase subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	547	Total	C	N	O	S	0	0	0
			4254	2848	666	710	30			
1	G	547	Total	C	N	O	S	0	0	0
			4292	2873	677	711	31			

- Molecule 2 is a protein called Cytochrome c oxidase subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	254	Total	C	N	O	S	0	0	0
			1956	1275	320	355	6			
2	H	253	Total	C	N	O	S	0	0	0
			1927	1265	316	340	6			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	282	HIS	-	expression tag	UNP Q03736
B	283	HIS	-	expression tag	UNP Q03736
B	284	HIS	-	expression tag	UNP Q03736
B	285	HIS	-	expression tag	UNP Q03736
B	286	HIS	-	expression tag	UNP Q03736
B	287	HIS	-	expression tag	UNP Q03736
H	282	HIS	-	expression tag	UNP Q03736
H	283	HIS	-	expression tag	UNP Q03736
H	284	HIS	-	expression tag	UNP Q03736
H	285	HIS	-	expression tag	UNP Q03736
H	286	HIS	-	expression tag	UNP Q03736
H	287	HIS	-	expression tag	UNP Q03736

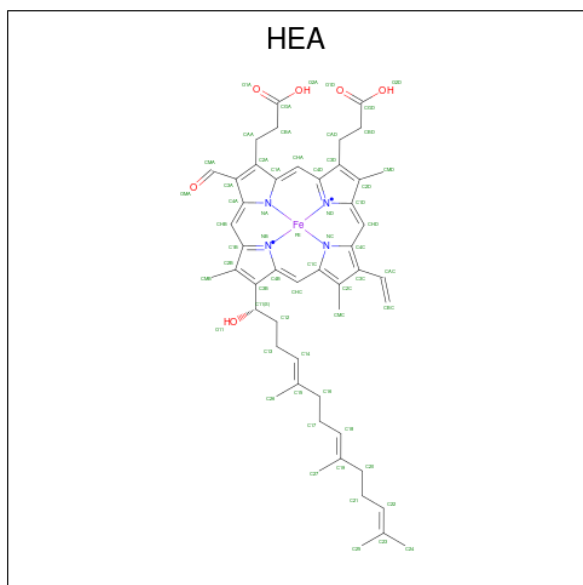
- Molecule 3 is a protein called Cytochrome c oxidase polypeptide III (Cytochrome AA3 subunit 3).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	264	Total	C	N	O	S	0	0	0
			2064	1387	331	334	12			
3	I	264	Total	C	N	O	S	0	0	0
			2095	1417	332	334	12			

- Molecule 4 is a protein called Aa3-type cytochrome c oxidase subunit IV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	40	Total	C	N	O	S	0	0	0
			276	183	43	49	1			
4	J	42	Total	C	N	O	S	0	0	0
			299	195	49	53	2			

- Molecule 5 is HEME-A (three-letter code: HEA) (formula:  $C_{49}H_{56}FeN_4O_6$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	Fe	N	O	0	0
			60	49	1	4	6		
5	A	1	Total	C	Fe	N	O	0	0
			60	49	1	4	6		
5	G	1	Total	C	Fe	N	O	0	0
			60	49	1	4	6		
5	G	1	Total	C	Fe	N	O	0	0
			60	49	1	4	6		

- Molecule 6 is COPPER (I) ION (three-letter code: CU1) (formula: Cu).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Cu 1 1	0	0
6	G	1	Total Cu 1 1	0	0
6	B	2	Total Cu 2 2	0	0
6	H	2	Total Cu 2 2	0	0

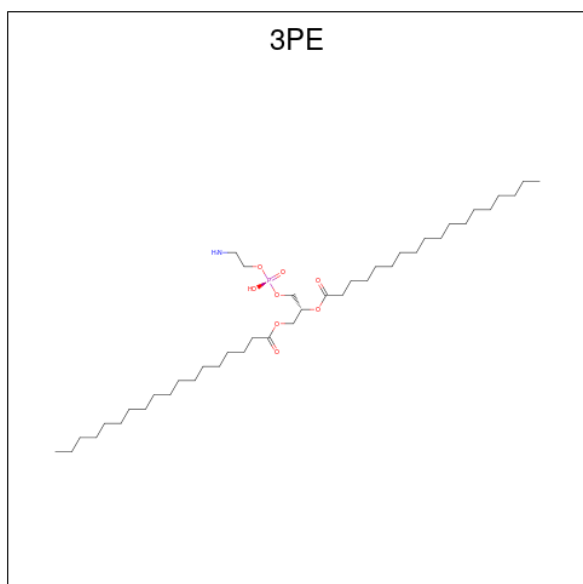
- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total Mg 1 1	0	0
7	G	1	Total Mg 1 1	0	0

- Molecule 8 is CALCIUM ION (three-letter code: CA) (formula: Ca).

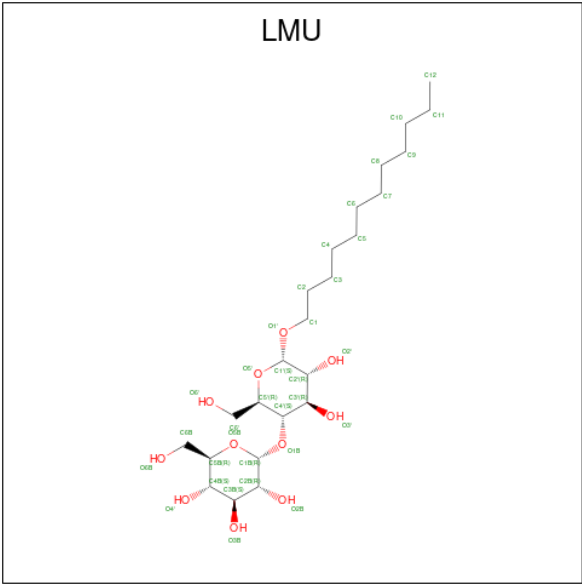
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total Ca 1 1	0	0
8	G	1	Total Ca 1 1	0	0

- Molecule 9 is 1,2-Distearoyl-sn-glycerophosphoethanolamine (three-letter code: 3PE) (formula: C<sub>41</sub>H<sub>82</sub>NO<sub>8</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	A	1	Total	C	N	O	P	0	0
			21	11	1	8	1		
9	A	1	Total	C	N	O	P	0	0
			45	35	1	8	1		
9	G	1	Total	C	N	O	P	0	0
			33	23	1	8	1		
9	G	1	Total	C	N	O	P	0	0
			44	34	1	8	1		
9	C	1	Total	C	N	O	P	0	0
			41	31	1	8	1		
9	C	1	Total	C	N	O	P	0	0
			51	41	1	8	1		
9	D	1	Total	C	N	O	P	0	0
			51	41	1	8	1		
9	I	1	Total	C	N	O	P	0	0
			40	30	1	8	1		
9	I	1	Total	C	N	O	P	0	0
			32	22	1	8	1		
9	I	1	Total	C	N	O	P	0	0
			23	13	1	8	1		
9	J	1	Total	C	N	O	P	0	0
			35	25	1	8	1		

- Molecule 10 is DODECYL-ALPHA-D-MALTOSIDE (three-letter code: LMU) (formula: C<sub>24</sub>H<sub>46</sub>O<sub>11</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	G	1	Total	C	O	0	0
			20	9	11		
10	G	1	Total	C	O	0	0
			24	13	11		
10	C	1	Total	C	O	0	0
			35	24	11		

- Molecule 11 is water.

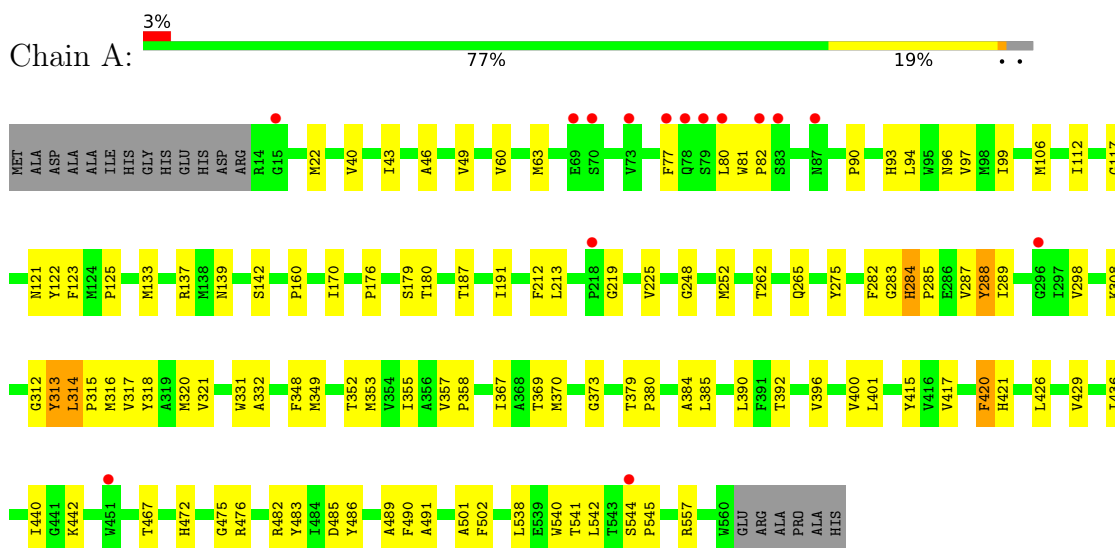
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	1	Total	O	0	0
			1	1		



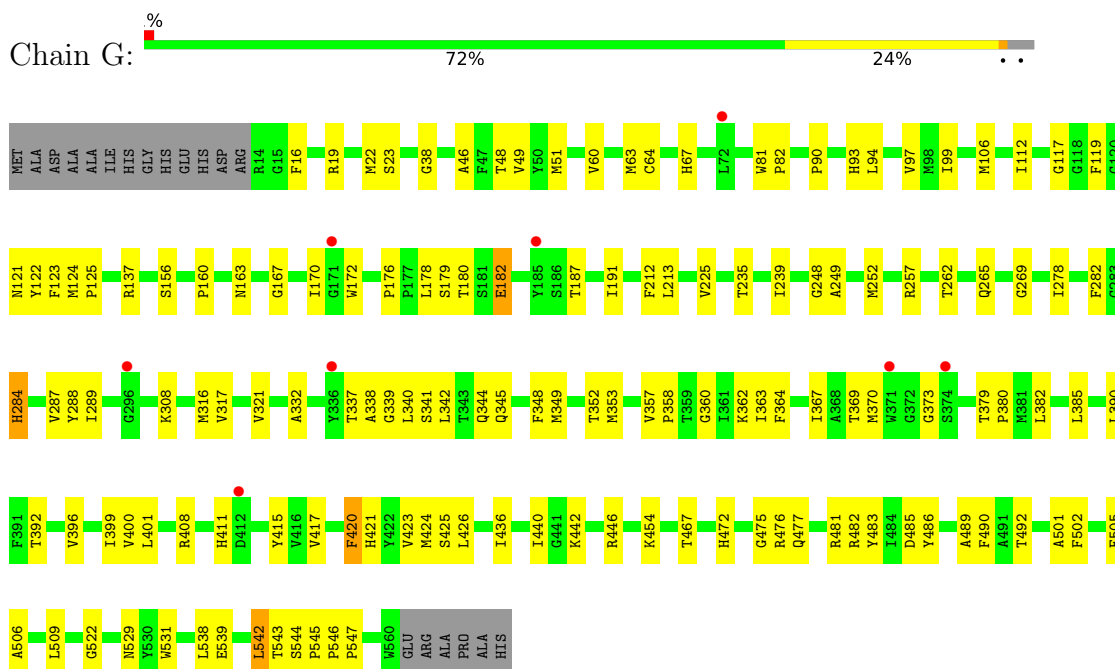
### 3 Residue-property plots [i](#)

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.

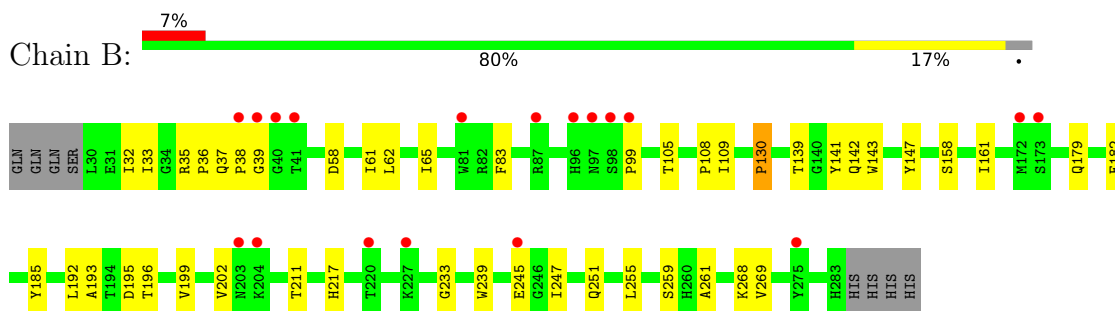
#### • Molecule 1: Cytochrome c oxidase subunit 1



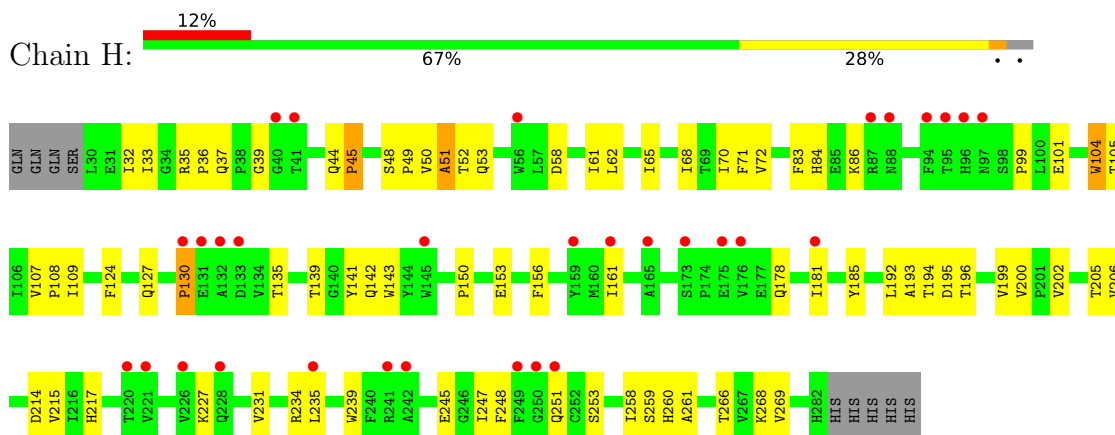
#### • Molecule 1: Cytochrome c oxidase subunit 1



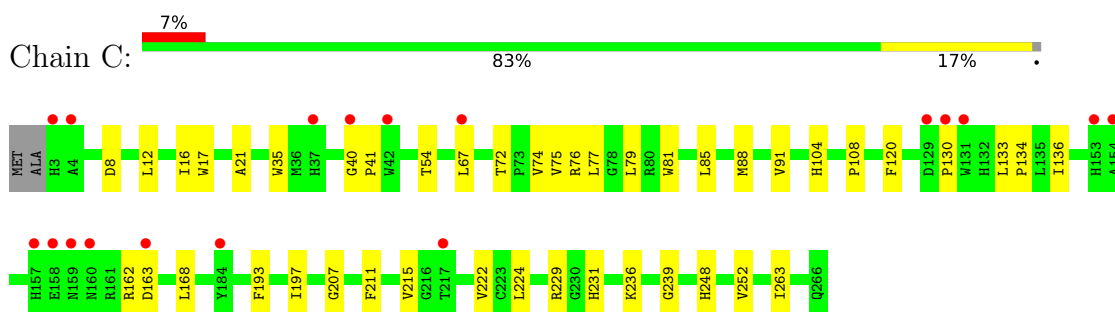
- Molecule 2: Cytochrome c oxidase subunit 2



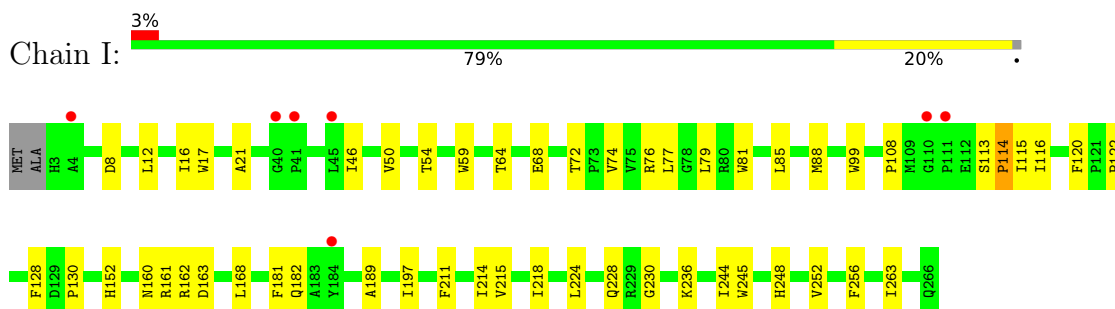
- Molecule 2: Cytochrome c oxidase subunit 2



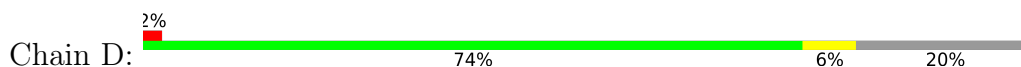
- Molecule 3: Cytochrome c oxidase polypeptide III (Cytochrome AA3 subunit 3)

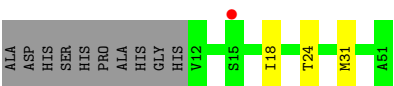


- Molecule 3: Cytochrome c oxidase polypeptide III (Cytochrome AA3 subunit 3)

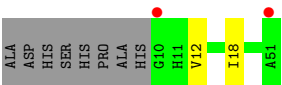
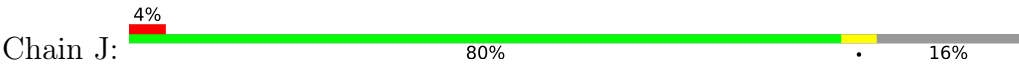


- Molecule 4: Aa3-type cytochrome c oxidase subunit IV





● Molecule 4: Aa3-type cytochrome c oxidase subunit IV



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	339.22Å 339.22Å 89.20Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.45 – 3.45 29.73 – 3.42	Depositor EDS
% Data completeness (in resolution range)	92.1 (28.45-3.45) 91.6 (29.73-3.42)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.73 (at 3.39Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, $R_{free}$	0.223 , 0.262 0.222 , 0.262	Depositor DCC
$R_{free}$ test set	1408 reflections (2.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	105.8	Xtriage
Anisotropy	0.759	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 79.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.54$ , $\langle L^2 \rangle = 0.38$	Xtriage
Estimated twinning fraction	0.000 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	17909	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% 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: CA, LMU, MG, 3PE, HEA, CU1

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	0.31	0/4413	0.49	3/6035 (0.0%)
1	G	0.31	1/4452 (0.0%)	0.49	2/6079 (0.0%)
2	B	0.27	0/2015	0.52	0/2765
2	H	0.32	0/1985	0.53	0/2726
3	C	0.27	0/2149	0.43	0/2943
3	I	0.29	0/2184	0.43	0/2991
4	D	0.25	0/280	0.39	0/384
4	J	0.26	0/303	0.41	0/411
All	All	0.30	1/17781 (0.0%)	0.48	5/24334 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	284	HIS	C-N	6.64	1.46	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	288	TYR	O-C-N	-5.83	113.36	122.70
1	G	288	TYR	O-C-N	-5.64	113.68	122.70
1	A	283	GLY	O-C-N	-5.34	114.16	122.70
1	G	288	TYR	C-N-CA	5.31	134.97	121.70
1	A	284	HIS	C-N-CD	5.12	139.16	128.40

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4254	0	4102	94	0
1	G	4292	0	4174	125	0
2	B	1956	0	1868	32	0
2	H	1927	0	1867	81	0
3	C	2064	0	1935	39	0
3	I	2095	0	1989	43	0
4	D	276	0	270	3	0
4	J	299	0	299	3	0
5	A	120	0	108	43	0
5	G	120	0	108	42	0
6	A	1	0	0	0	0
6	B	2	0	0	0	0
6	G	1	0	0	0	0
6	H	2	0	0	0	0
7	A	1	0	0	0	0
7	G	1	0	0	0	0
8	A	1	0	0	0	0
8	G	1	0	0	0	0
9	A	66	0	83	6	0
9	C	92	0	141	10	0
9	D	51	0	82	1	0
9	G	77	0	105	2	0
9	I	95	0	112	3	0
9	J	35	0	47	1	0
10	C	35	0	46	2	0
10	G	44	0	28	15	0
11	A	1	0	0	0	0
All	All	17909	0	17364	445	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 13.

The worst 5 of 445 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)
5:A:602:HEA:CHA	5:A:602:HEA:C1A	1.75	1.63
5:G:602:HEA:C4B	5:G:602:HEA:CHC	1.80	1.59
5:G:603:HEA:CHC	5:G:603:HEA:C4B	1.80	1.59
5:G:603:HEA:CHD	5:G:603:HEA:C1D	1.81	1.59
5:A:601:HEA:CHC	5:A:601:HEA:C4B	1.79	1.58

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	545/566 (96%)	502 (92%)	40 (7%)	3 (1%)	25	62
1	G	545/566 (96%)	504 (92%)	39 (7%)	2 (0%)	34	70
2	B	252/262 (96%)	227 (90%)	23 (9%)	2 (1%)	19	57
2	H	251/262 (96%)	220 (88%)	26 (10%)	5 (2%)	7	37
3	C	262/266 (98%)	247 (94%)	15 (6%)	0	100	100
3	I	262/266 (98%)	246 (94%)	15 (6%)	1 (0%)	34	70
4	D	38/50 (76%)	37 (97%)	1 (3%)	0	100	100
4	J	40/50 (80%)	39 (98%)	1 (2%)	0	100	100
All	All	2195/2288 (96%)	2022 (92%)	160 (7%)	13 (1%)	25	62

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	45	PRO
2	H	153	GLU
1	G	16	PHE
2	H	51	ALA
1	A	313	TYR

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	431/459 (94%)	426 (99%)	5 (1%)	71	87
1	G	439/459 (96%)	433 (99%)	6 (1%)	67	85
2	B	199/221 (90%)	199 (100%)	0	100	100
2	H	195/221 (88%)	194 (100%)	1 (0%)	88	95
3	C	200/216 (93%)	200 (100%)	0	100	100
3	I	207/216 (96%)	203 (98%)	4 (2%)	57	80
4	D	24/36 (67%)	24 (100%)	0	100	100
4	J	27/36 (75%)	27 (100%)	0	100	100
All	All	1722/1864 (92%)	1706 (99%)	16 (1%)	78	91

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	182	GLN
3	I	181	PHE
1	G	485	ASP
3	I	114	PRO
1	G	420	PHE

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

Mol	Chain	Res	Type
1	A	334	HIS
2	B	217	HIS
2	H	217	HIS
2	H	260	HIS
3	I	157	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.



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

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

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 28 ligands modelled in this entry, 10 are monoatomic - leaving 18 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$
10	LMU	G	609	-	25,25,36	1.28	2 (8%)	36,36,47	0.97	2 (5%)
9	3PE	C	301	-	40,40,50	1.01	3 (7%)	43,45,55	1.09	2 (4%)
9	3PE	A	606	-	20,20,50	1.32	4 (20%)	23,25,55	1.18	2 (8%)
9	3PE	J	101	-	34,34,50	0.98	3 (8%)	37,39,55	1.02	1 (2%)
9	3PE	G	608	-	43,43,50	0.93	3 (6%)	46,48,55	1.08	2 (4%)
9	3PE	I	301	-	39,39,50	1.00	4 (10%)	42,44,55	1.11	2 (4%)
9	3PE	A	607	-	44,44,50	0.92	4 (9%)	47,49,55	1.10	3 (6%)
9	3PE	D	101	-	50,50,50	0.90	3 (6%)	53,55,55	1.06	2 (3%)
5	HEA	G	602	1	44,67,67	4.79	17 (38%)	37,103,103	3.32	22 (59%)
10	LMU	G	601	-	13,13,36	0.63	0	10,10,47	0.55	0
9	3PE	C	302	-	50,50,50	0.88	4 (8%)	53,55,55	1.07	2 (3%)
9	3PE	G	607	-	32,32,50	1.06	3 (9%)	35,37,55	1.10	2 (5%)
9	3PE	I	302	-	31,31,50	1.08	4 (12%)	34,36,55	1.16	2 (5%)
10	LMU	C	303	-	36,36,36	1.23	3 (8%)	47,47,47	1.48	5 (10%)
5	HEA	A	601	1	44,67,67	4.71	16 (36%)	37,103,103	3.15	18 (48%)
5	HEA	A	602	1	44,67,67	5.06	15 (34%)	37,103,103	3.07	14 (37%)
5	HEA	G	603	1	44,67,67	5.06	18 (40%)	37,103,103	2.75	17 (45%)
9	3PE	I	303	-	22,22,50	1.27	4 (18%)	25,27,55	1.13	2 (8%)

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
10	LMU	G	609	-	-	4/10/50/61	0/2/2/2
9	3PE	C	301	-	-	18/44/44/54	-
9	3PE	A	606	-	-	11/24/24/54	-
9	3PE	J	101	-	-	21/37/37/54	-
9	3PE	G	608	-	-	25/47/47/54	-
9	3PE	I	301	-	-	24/43/43/54	-
5	HEA	G	603	1	1/1/7/16	4/24/76/76	-
9	3PE	A	607	-	-	25/48/48/54	-
5	HEA	G	602	1	-	3/24/76/76	-
9	3PE	D	101	-	-	33/54/54/54	-
10	LMU	G	601	-	-	4/8/8/61	-
9	3PE	C	302	-	-	30/54/54/54	-
9	3PE	G	607	-	-	12/36/36/54	-
9	3PE	I	302	-	-	15/35/35/54	-
10	LMU	C	303	-	-	12/21/61/61	0/2/2/2
5	HEA	A	602	1	1/1/7/16	6/24/76/76	-
5	HEA	A	601	1	-	7/24/76/76	-
9	3PE	I	303	-	-	17/26/26/54	-

The worst 5 of 110 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	602	HEA	C1C-CHC	17.92	1.90	1.41
5	G	603	HEA	C1C-CHC	17.14	1.88	1.41
5	G	602	HEA	C4C-CHD	16.81	1.87	1.41
5	G	603	HEA	C4C-CHD	16.60	1.87	1.41
5	A	601	HEA	C4C-CHD	16.41	1.86	1.41

The worst 5 of 100 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	G	602	HEA	CAD-CBD-CGD	-9.48	96.76	112.67
5	A	602	HEA	C3C-C4C-NC	8.94	120.76	109.21
5	A	601	HEA	CAD-CBD-CGD	-8.77	97.96	112.67

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	601	HEA	C3C-C4C-NC	7.83	119.34	109.21
5	G	602	HEA	C3C-C4C-NC	7.81	119.31	109.21

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	602	HEA	NA
5	G	603	HEA	NA

5 of 271 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	601	HEA	C1A-C2A-CAA-CBA
5	A	601	HEA	C3A-C2A-CAA-CBA
5	A	601	HEA	C3B-C11-C12-C13
5	A	601	HEA	O11-C11-C12-C13
5	A	601	HEA	C11-C12-C13-C14

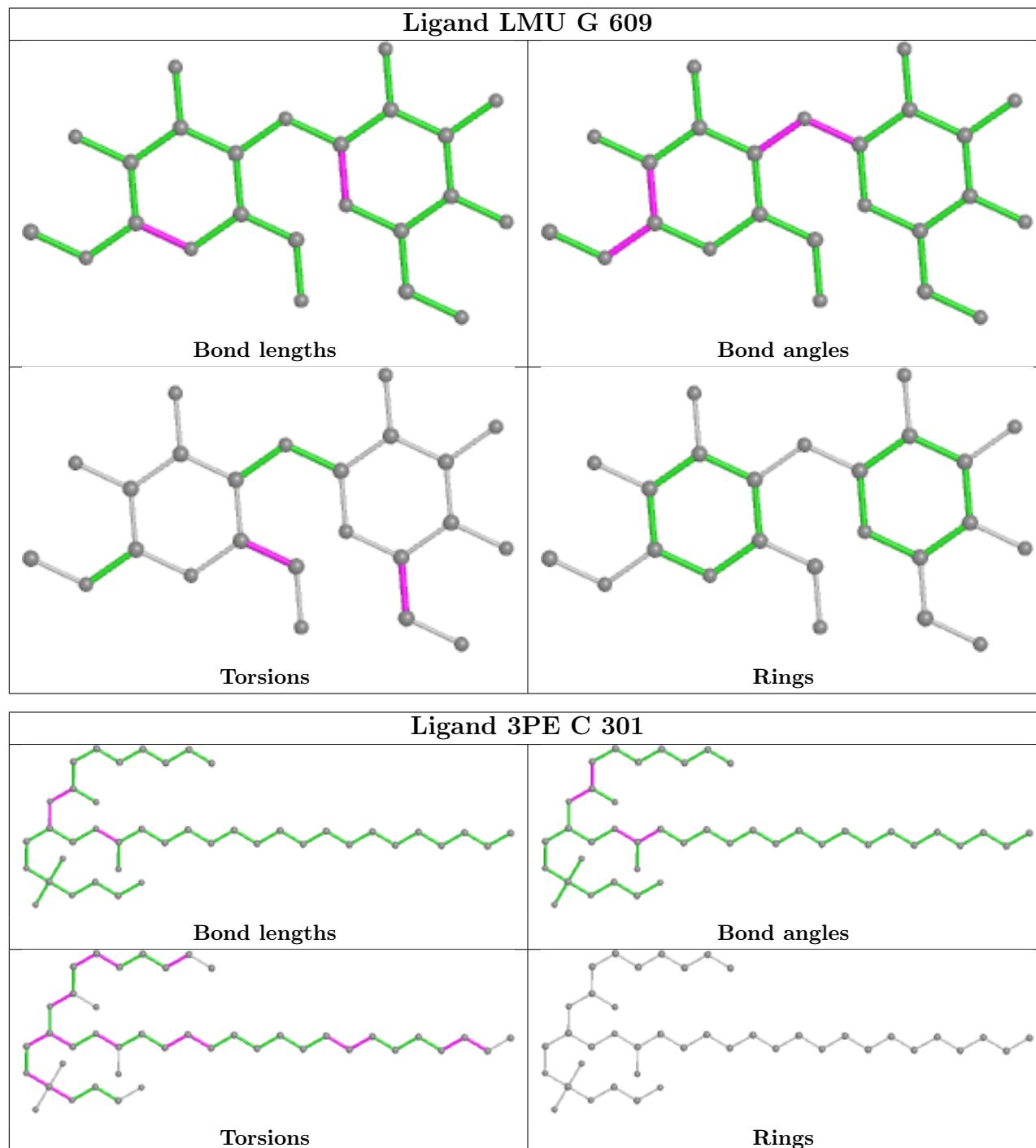
There are no ring outliers.

15 monomers are involved in 122 short contacts:

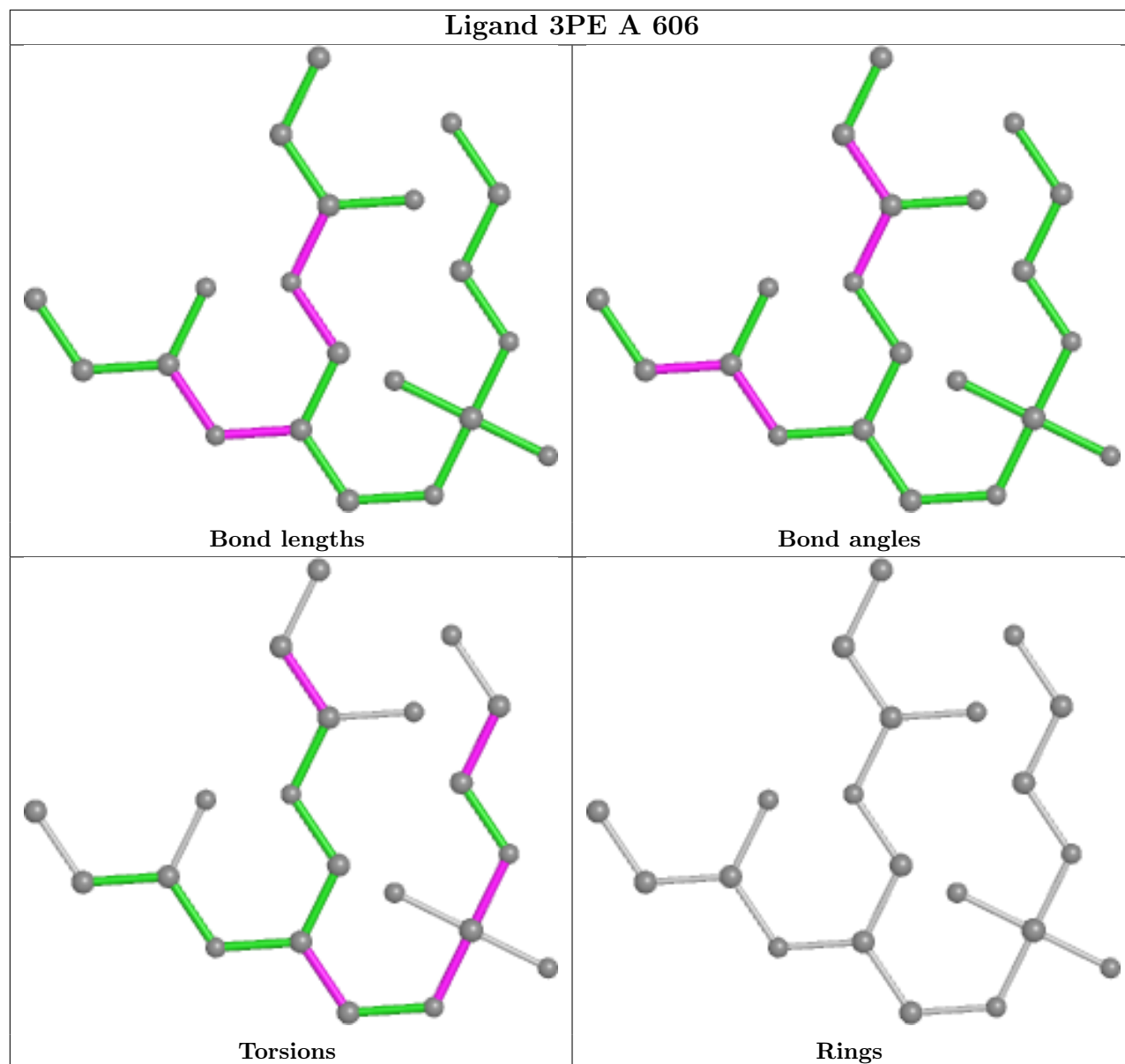
Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	C	301	3PE	4	0
9	A	606	3PE	2	0
9	J	101	3PE	1	0
9	G	608	3PE	2	0
9	I	301	3PE	2	0
9	A	607	3PE	4	0
9	D	101	3PE	1	0
5	G	602	HEA	23	0
10	G	601	LMU	15	0
9	C	302	3PE	6	0
10	C	303	LMU	2	0
5	A	601	HEA	19	0
5	A	602	HEA	24	0
5	G	603	HEA	19	0
9	I	303	3PE	1	0

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

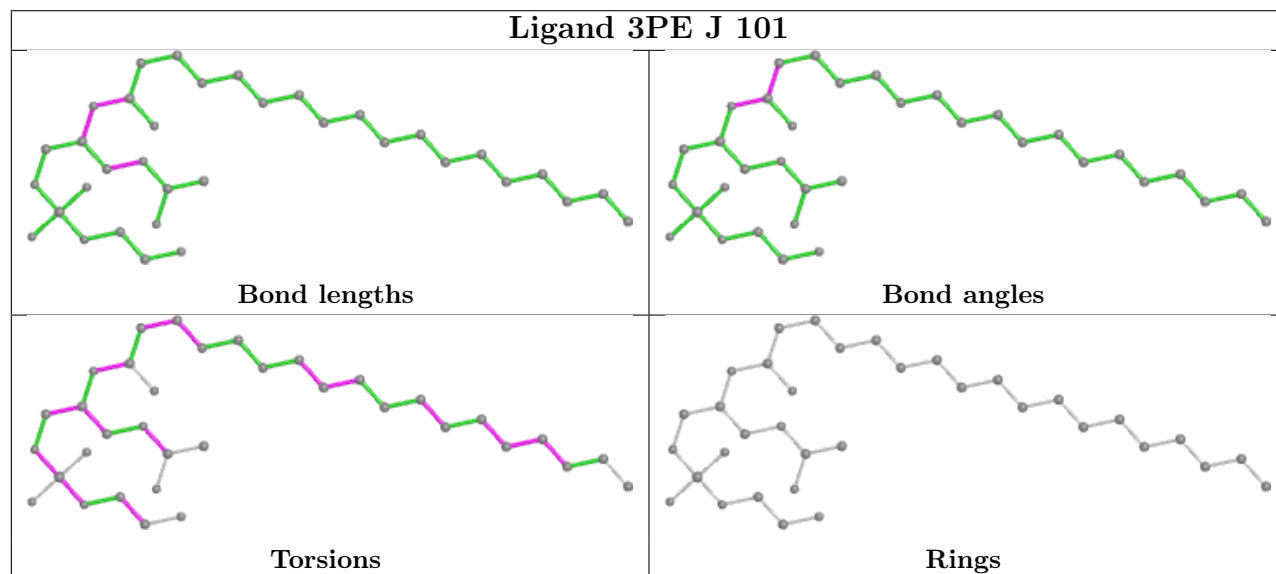
within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question 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.

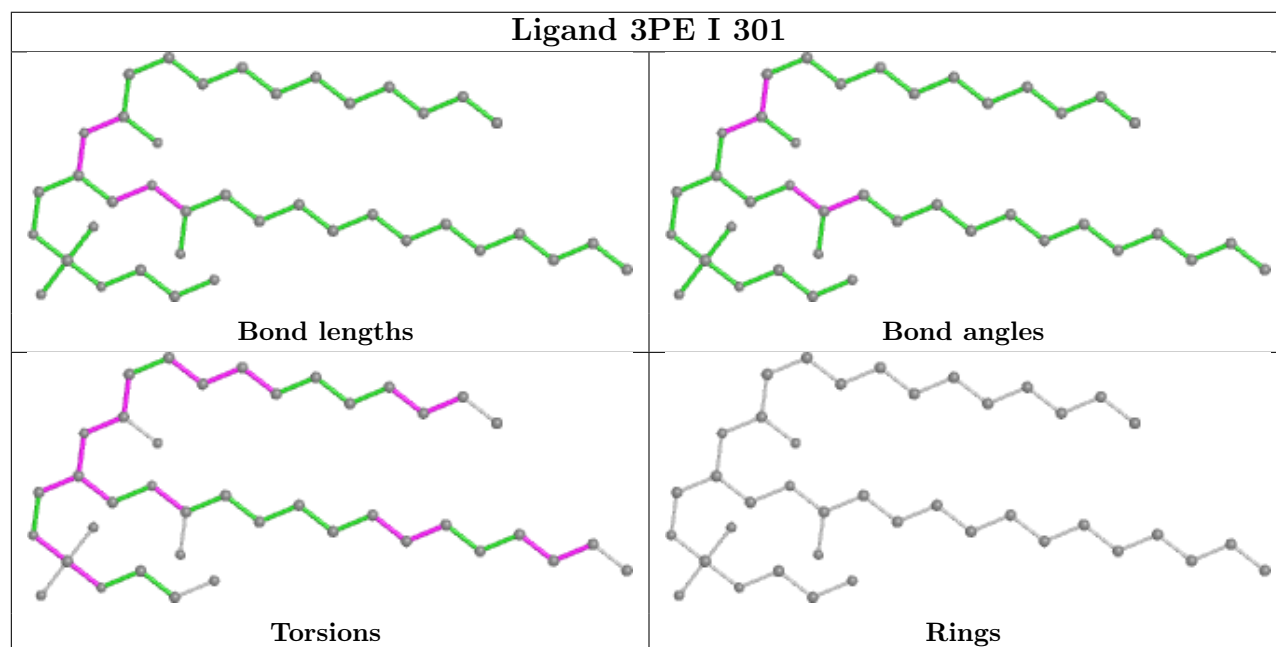
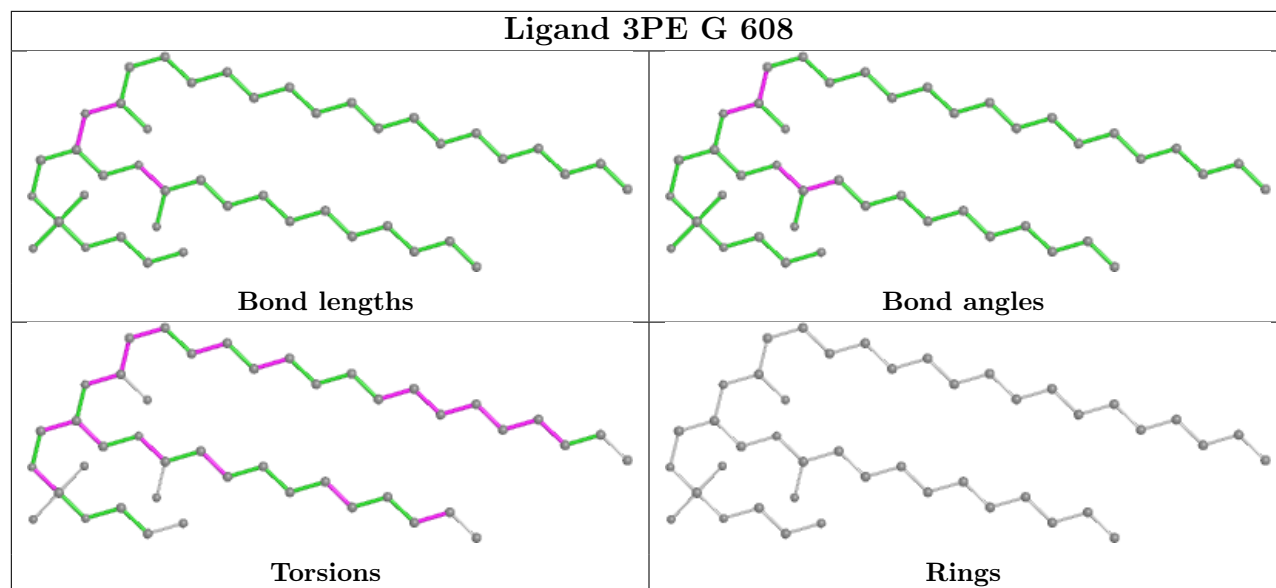


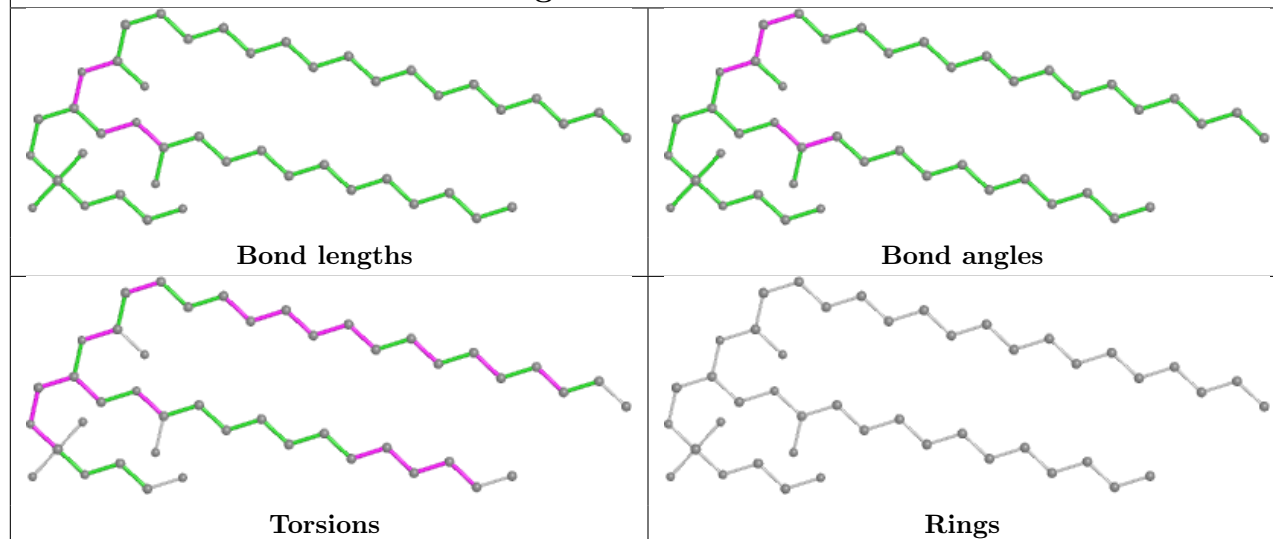
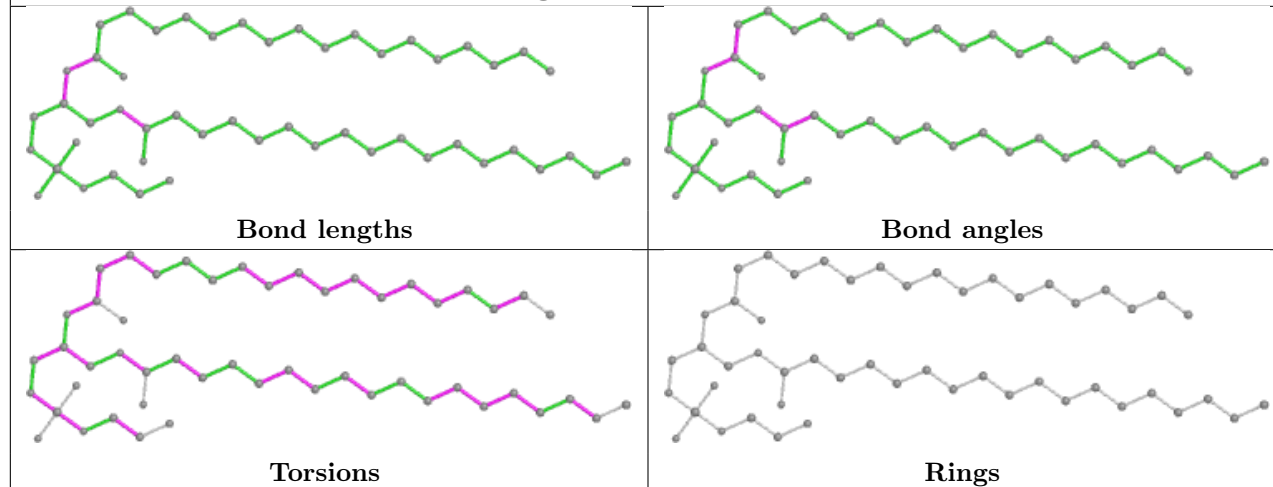
## Ligand 3PE A 606

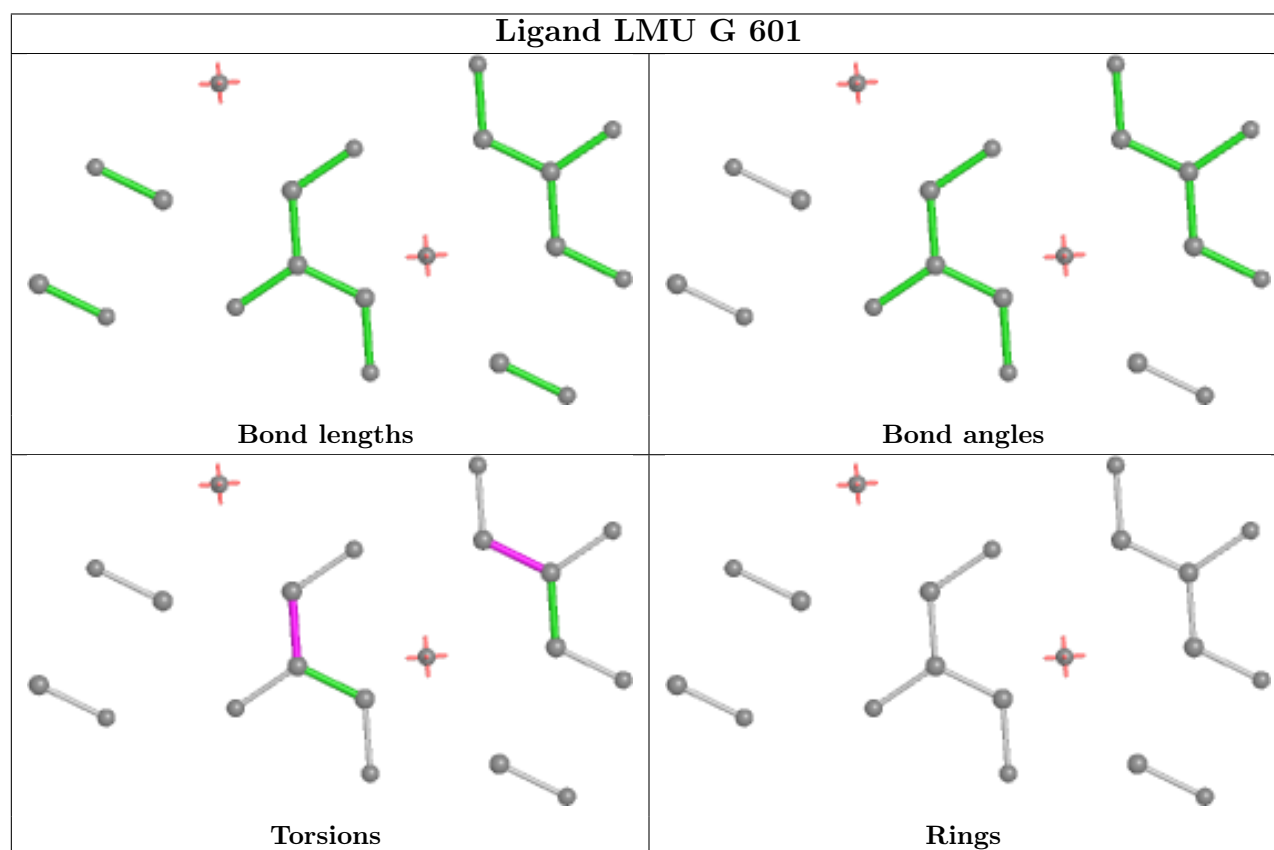
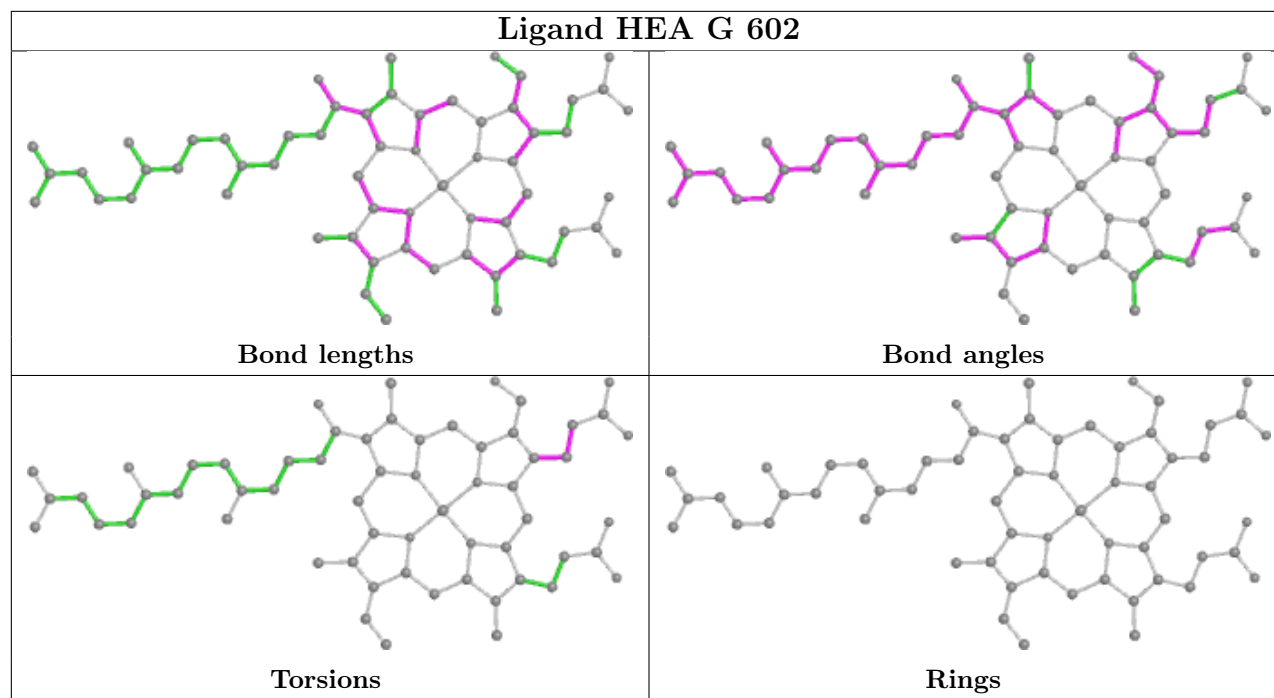


## Ligand 3PE J 101

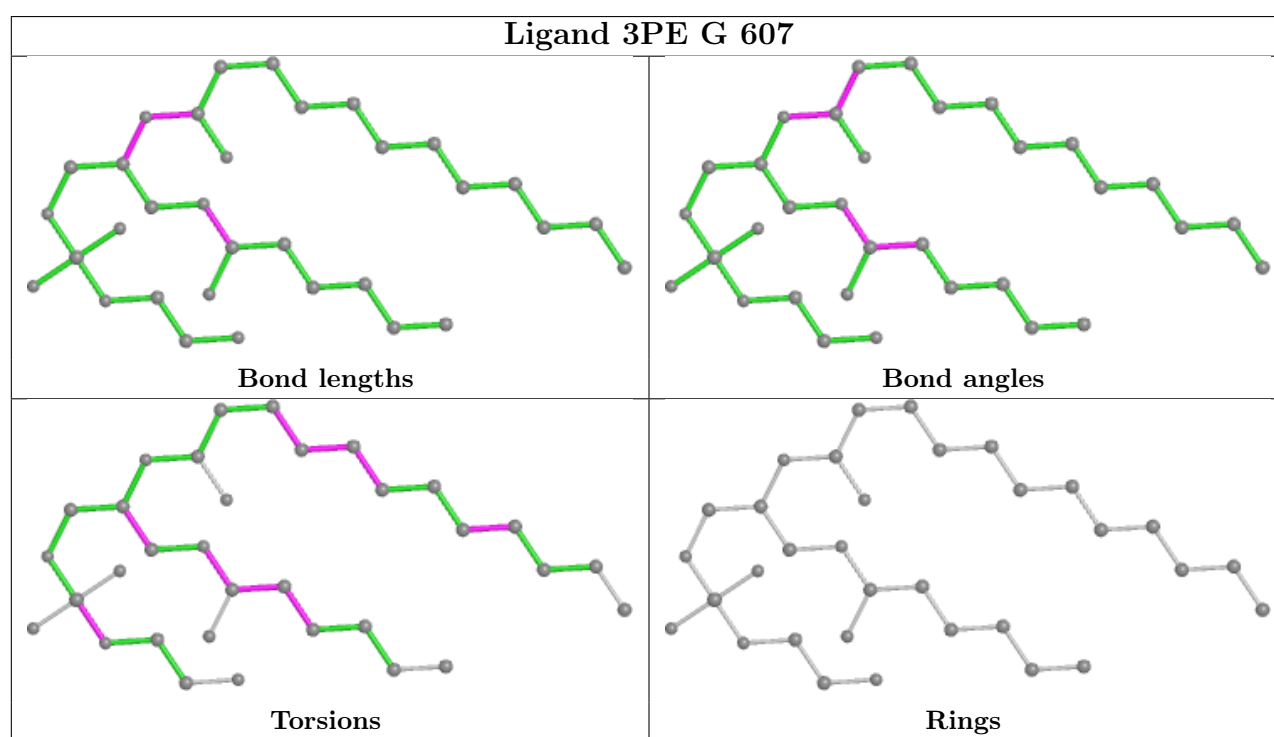
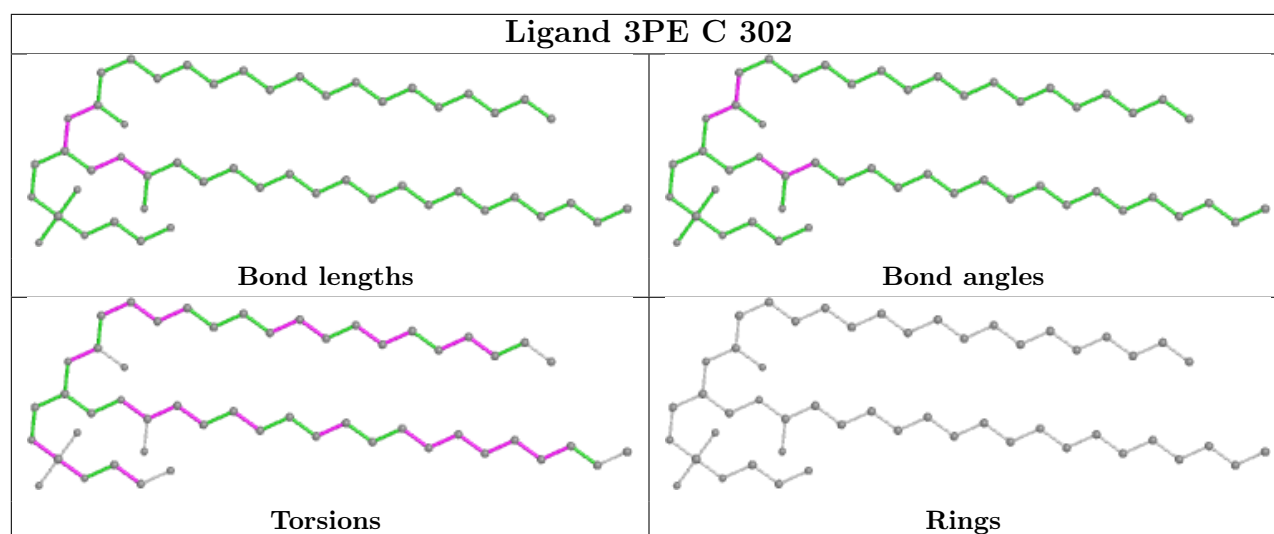


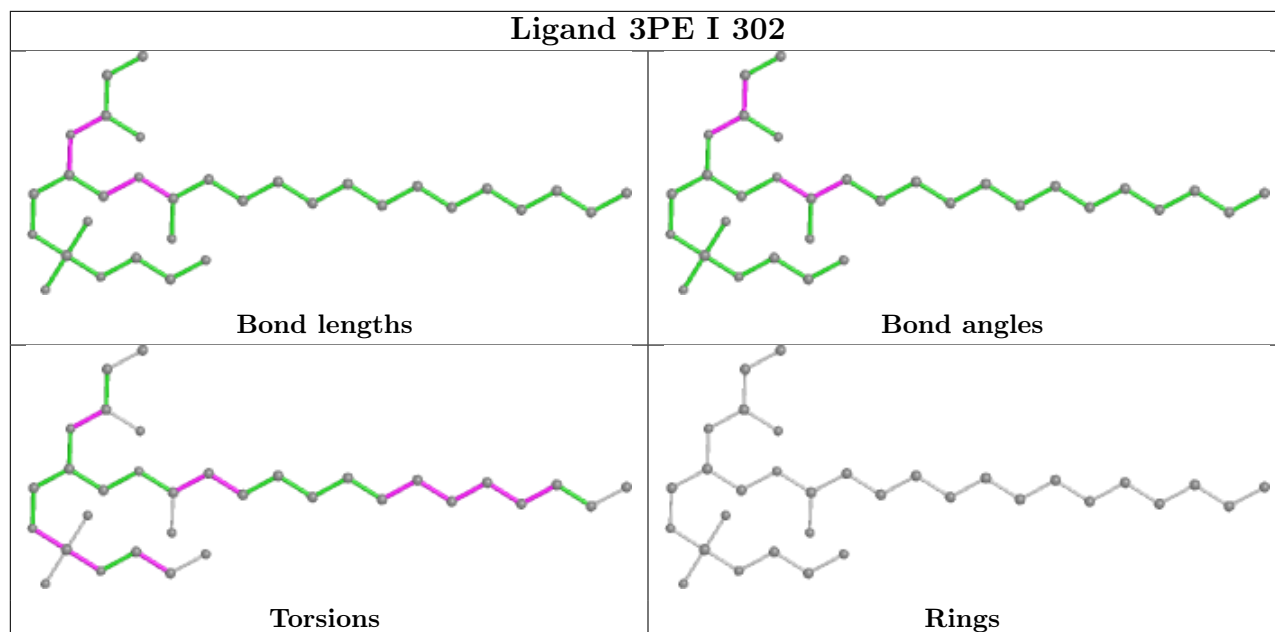
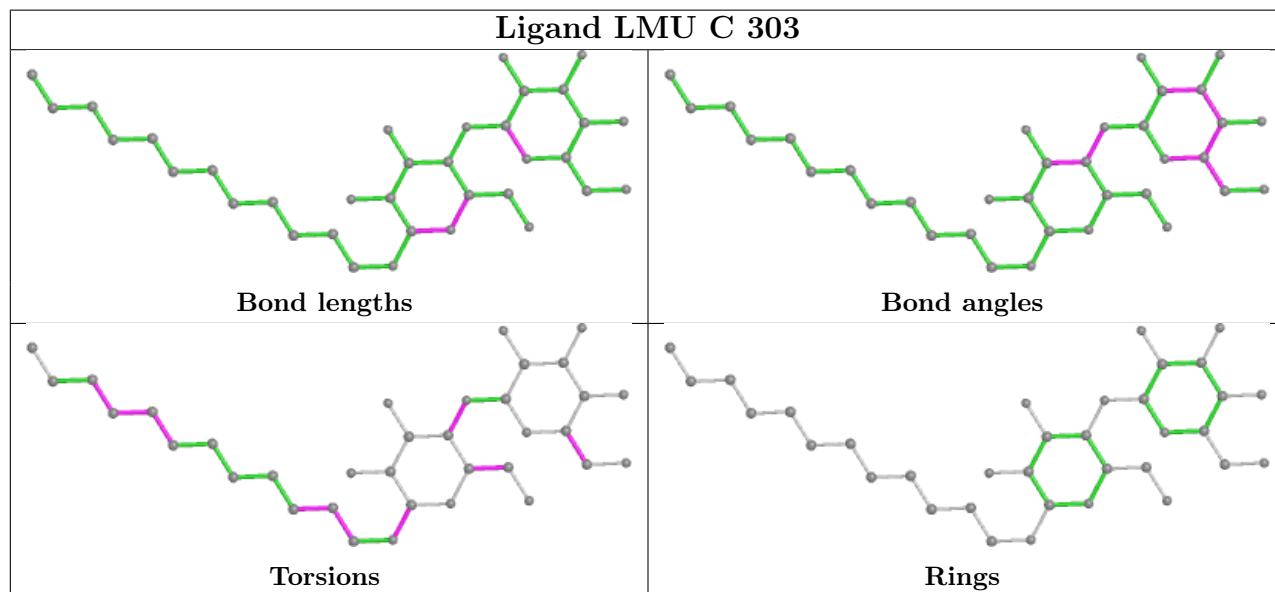


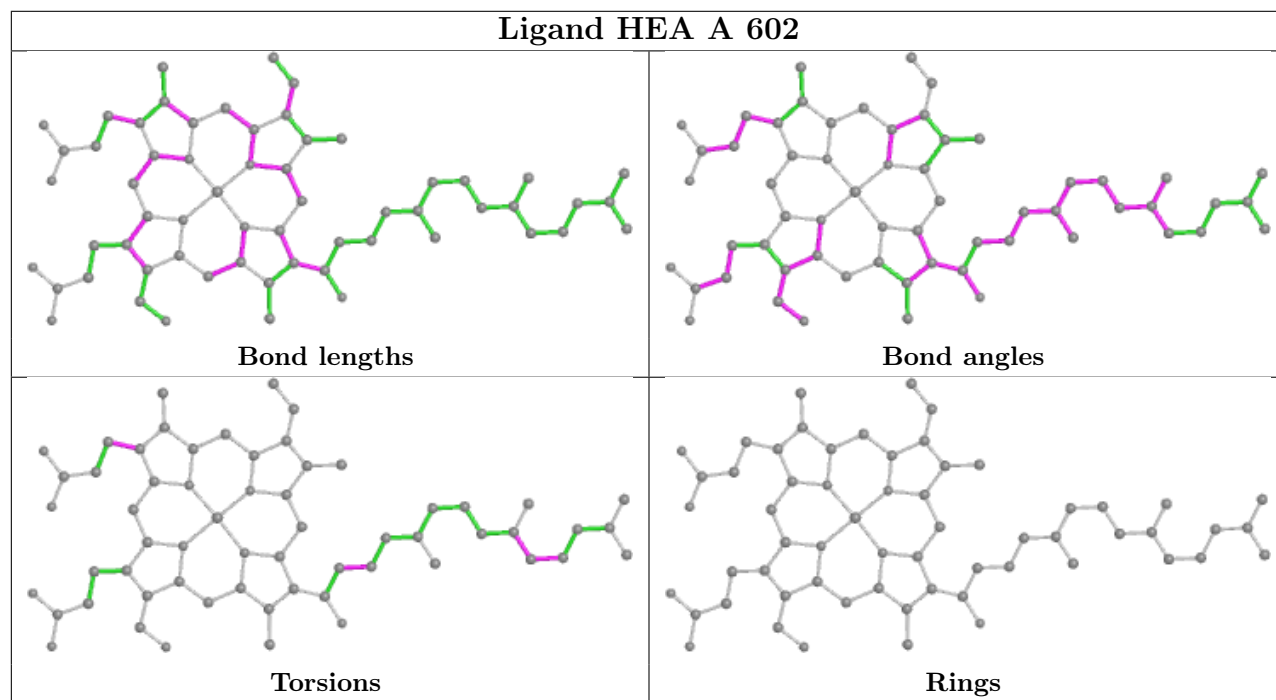
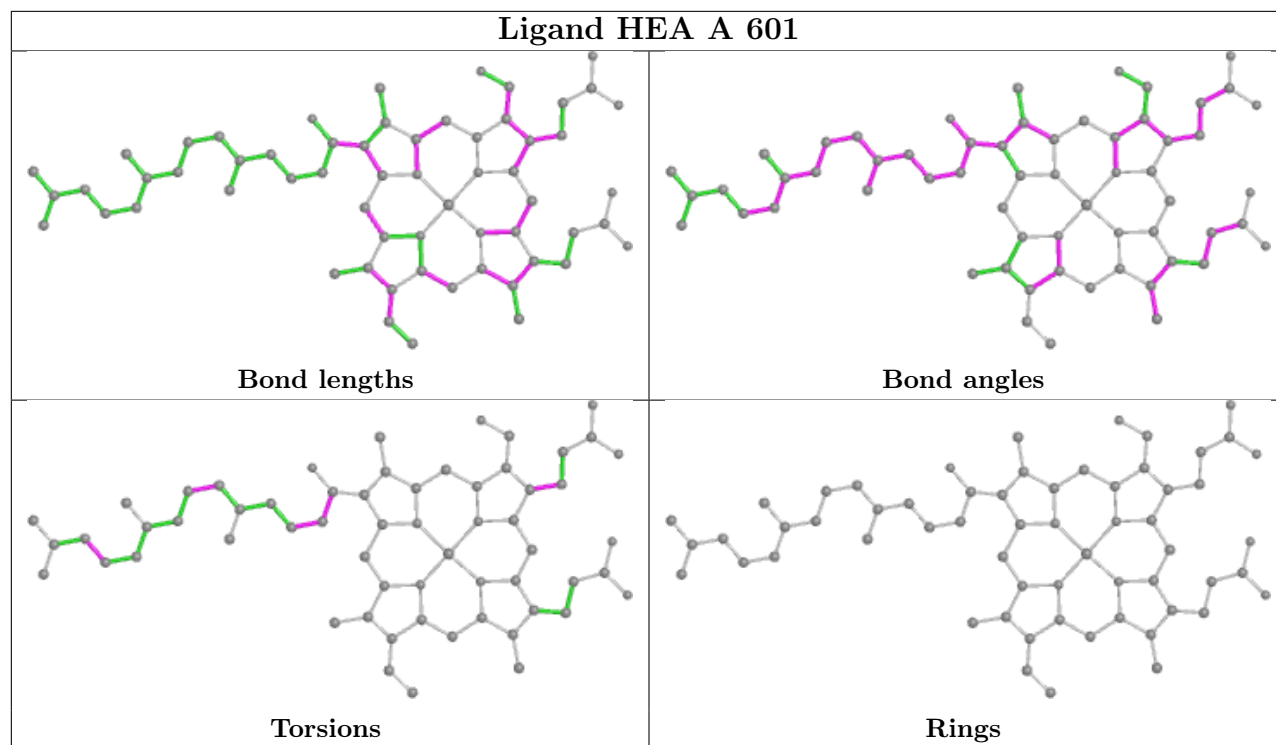
**Ligand 3PE A 607****Ligand 3PE D 101**

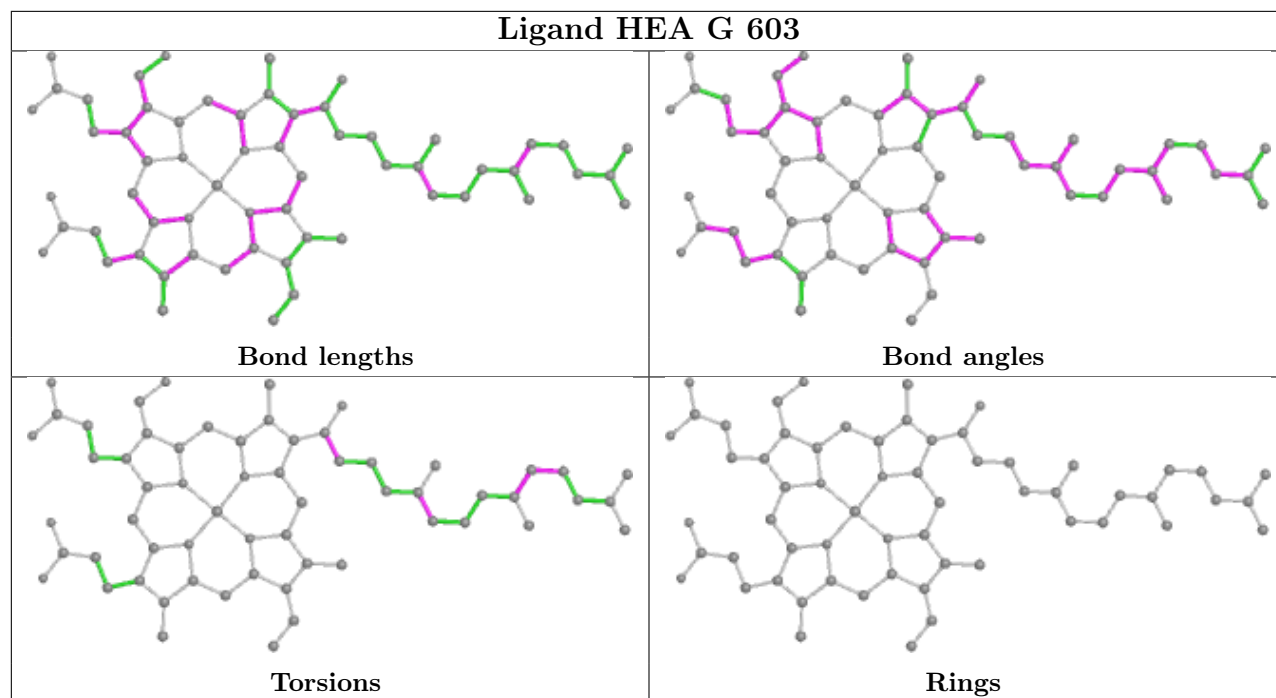


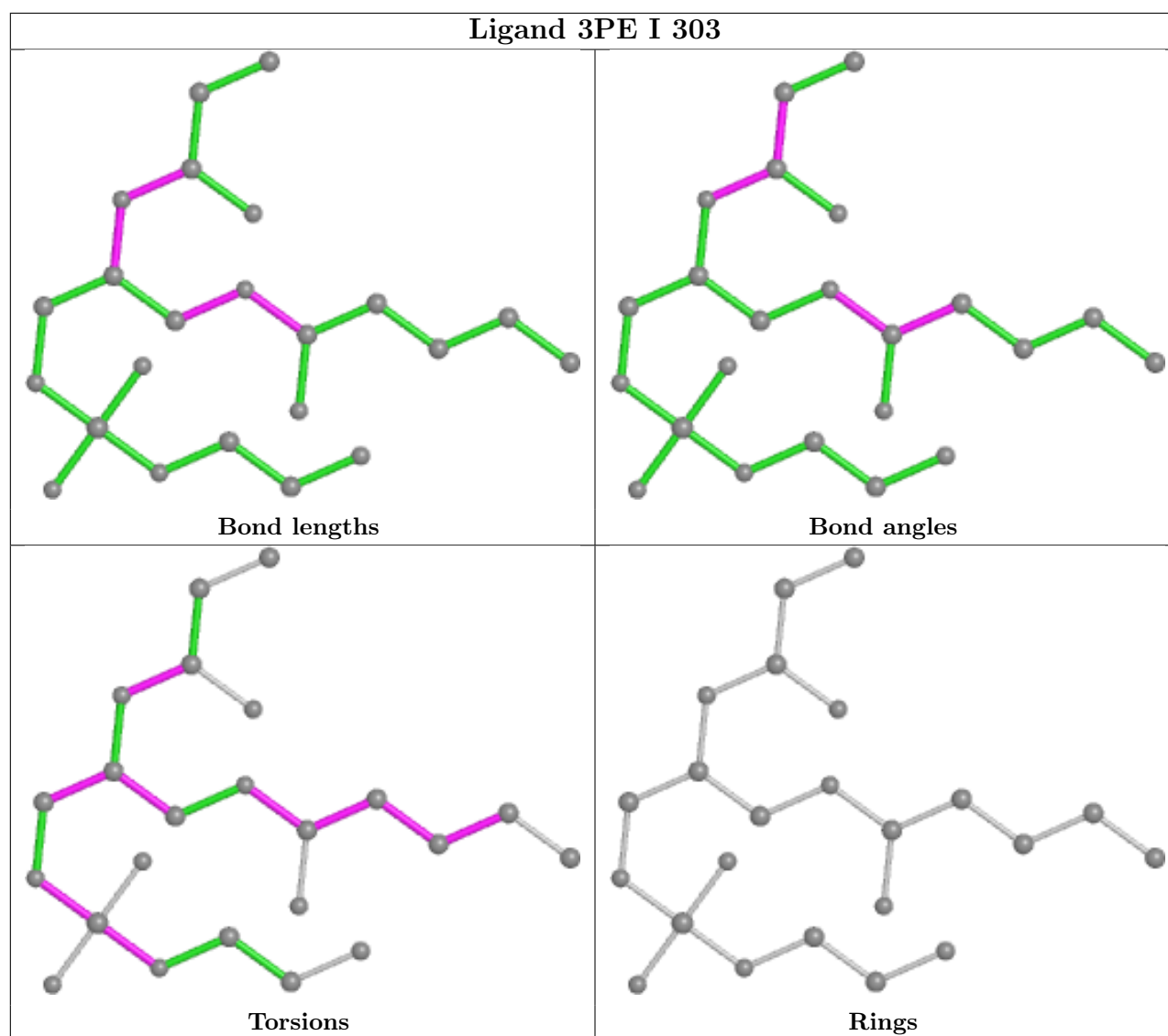




**Ligand 3PE I 302****Ligand LMU C 303**







## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

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

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

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	547/566 (96%)	0.05	15 (2%) 54 52	76, 103, 142, 188	0
1	G	547/566 (96%)	-0.02	8 (1%) 73 71	75, 103, 136, 180	0
2	B	254/262 (96%)	0.31	18 (7%) 16 18	83, 118, 152, 165	0
2	H	253/262 (96%)	0.59	31 (12%) 4 6	95, 135, 168, 178	0
3	C	264/266 (99%)	0.42	18 (6%) 17 19	92, 135, 168, 185	0
3	I	264/266 (99%)	-0.02	7 (2%) 54 52	79, 115, 151, 179	0
4	D	40/50 (80%)	0.14	1 (2%) 57 54	90, 120, 148, 160	0
4	J	42/50 (84%)	-0.17	2 (4%) 30 30	88, 114, 143, 166	0
All	All	2211/2288 (96%)	0.16	100 (4%) 33 32	75, 113, 157, 188	0

The worst 5 of 100 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	42	TRP	5.1
2	H	250	GLY	5.0
3	C	163	ASP	4.7
2	H	132	ALA	4.3
3	C	159	ASN	4.0

### 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 ⓘ

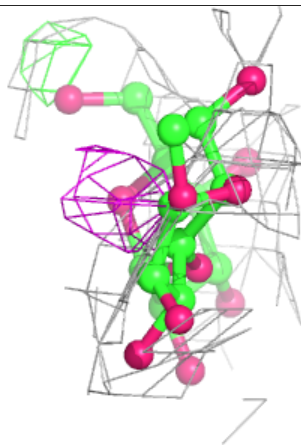
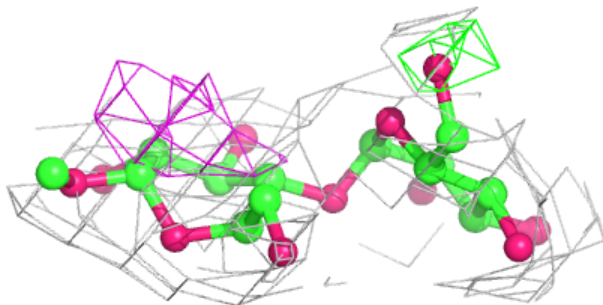
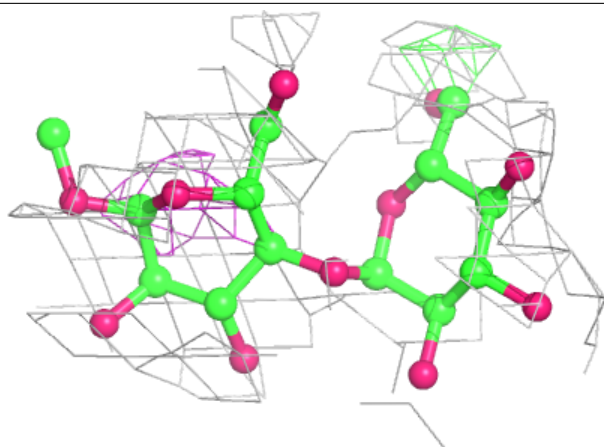
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
10	LMU	G	609	24/35	0.51	0.44	131,164,170,176	0
9	3PE	I	303	23/51	0.56	0.43	101,155,183,196	0
10	LMU	G	601	20/35	0.62	0.31	84,115,124,143	5
7	MG	G	605	1/1	0.66	0.58	86,86,86,86	0
9	3PE	C	301	41/51	0.74	0.48	100,157,200,216	0
9	3PE	I	302	32/51	0.78	0.34	86,124,183,198	0
10	LMU	C	303	35/35	0.79	0.38	108,150,203,214	0
8	CA	A	605	1/1	0.80	0.13	102,102,102,102	0
9	3PE	C	302	51/51	0.81	0.36	92,111,150,185	0
9	3PE	A	607	45/51	0.82	0.48	65,93,109,138	45
6	CU1	G	604	1/1	0.82	0.11	94,94,94,94	0
9	3PE	G	608	44/51	0.84	0.39	78,110,175,179	0
9	3PE	A	606	21/51	0.85	0.32	97,112,119,134	21
9	3PE	I	301	40/51	0.86	0.34	88,112,150,154	0
7	MG	A	604	1/1	0.86	0.58	95,95,95,95	0
9	3PE	J	101	35/51	0.89	0.38	97,123,164,172	0
9	3PE	D	101	51/51	0.89	0.31	74,105,145,150	0
9	3PE	G	607	33/51	0.90	0.33	70,82,98,107	0
6	CU1	H	302	1/1	0.90	0.22	107,107,107,107	0
8	CA	G	606	1/1	0.92	0.10	103,103,103,103	0
5	HEA	G	603	60/60	0.92	0.26	94,105,122,134	0
5	HEA	A	602	60/60	0.92	0.29	87,97,105,113	0
5	HEA	G	602	60/60	0.93	0.30	83,95,104,106	0
5	HEA	A	601	60/60	0.93	0.30	86,98,107,127	0
6	CU1	A	603	1/1	0.94	0.14	81,81,81,81	0
6	CU1	B	301	1/1	0.96	0.26	105,105,105,105	0
6	CU1	H	301	1/1	0.98	0.16	117,117,117,117	0
6	CU1	B	302	1/1	0.98	0.17	102,102,102,102	0

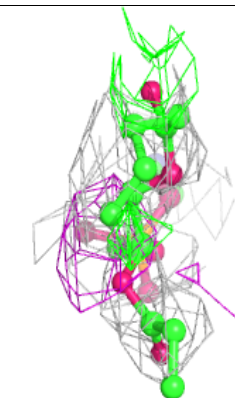
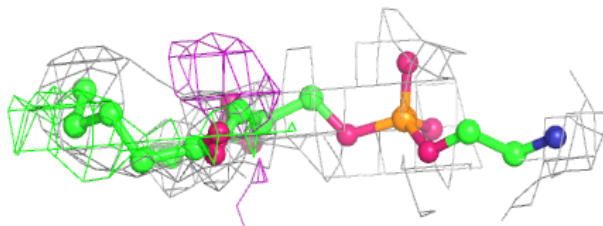
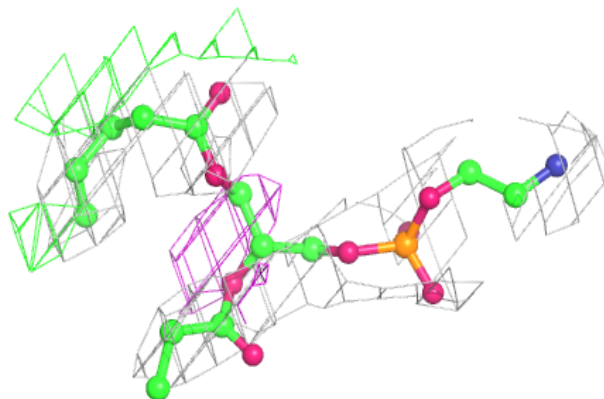
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around LMU G 609:**

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

**Electron density around 3PE I 303:**

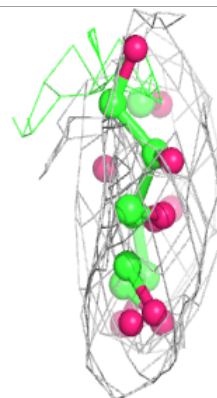
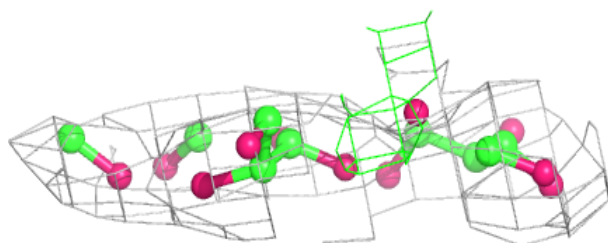
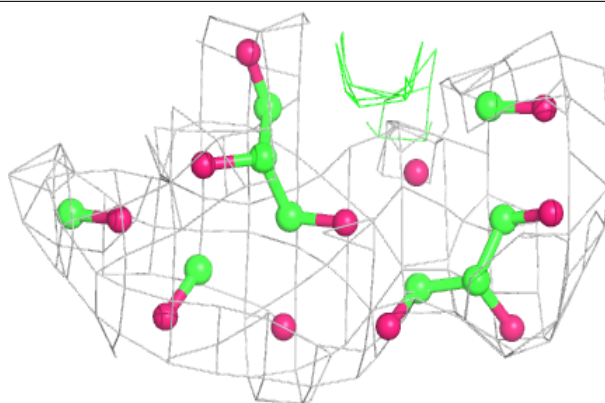
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



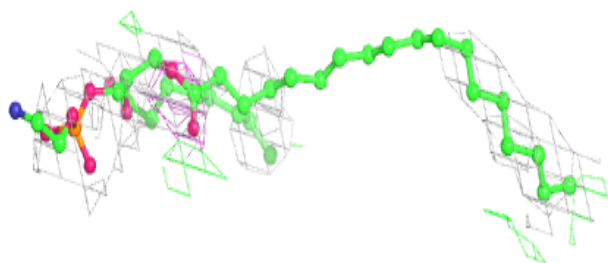
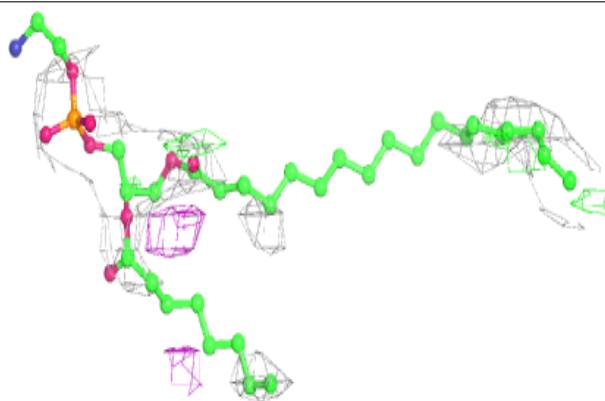


**Electron density around LMU G 601:**

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

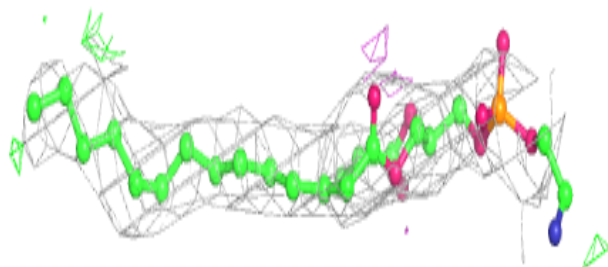
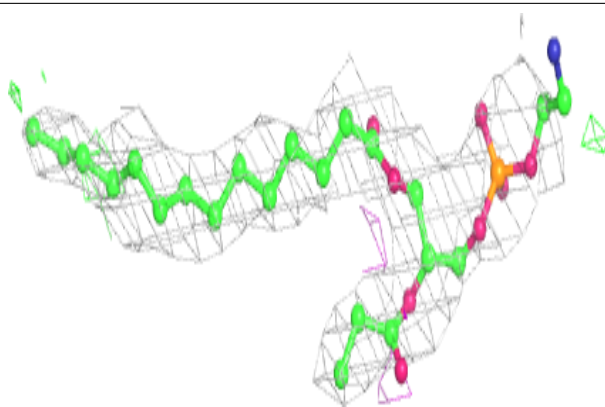
**Electron density around 3PE C 301:**

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

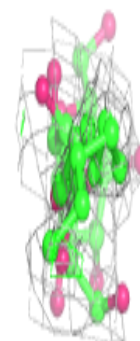
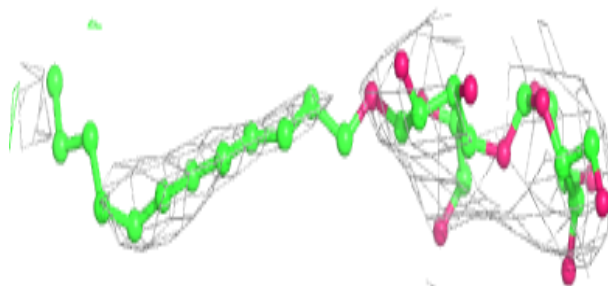
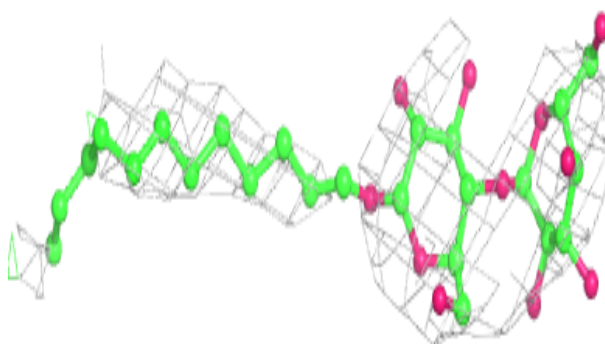


**Electron density around 3PE I 302:**

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

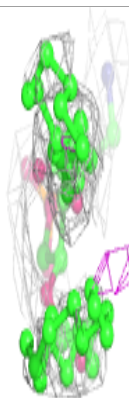
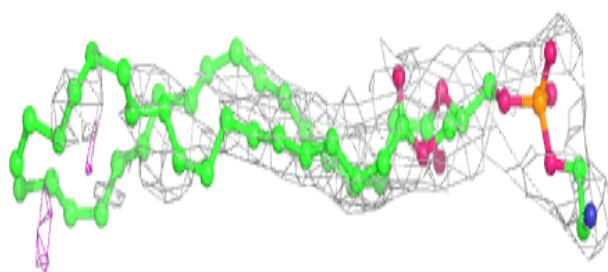
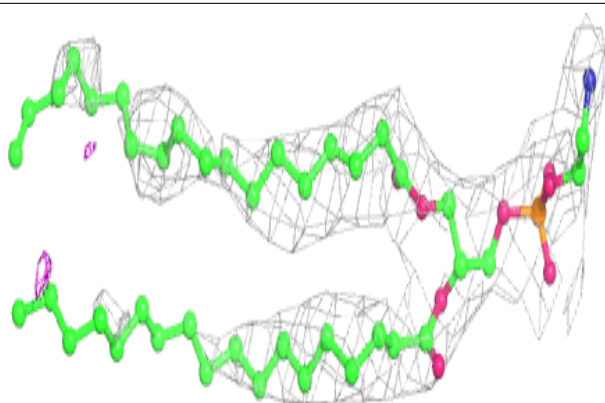
**Electron density around LMU C 303:**

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

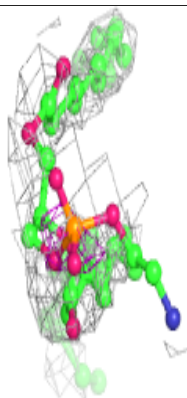
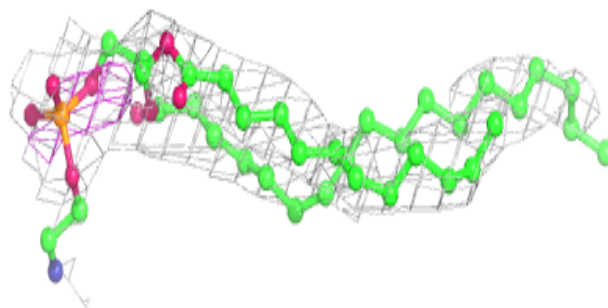
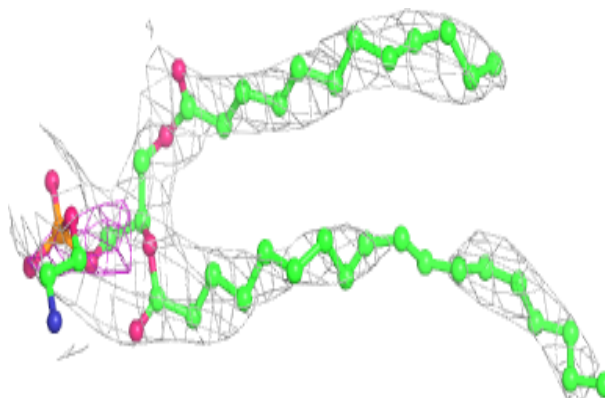


**Electron density around 3PE C 302:**

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

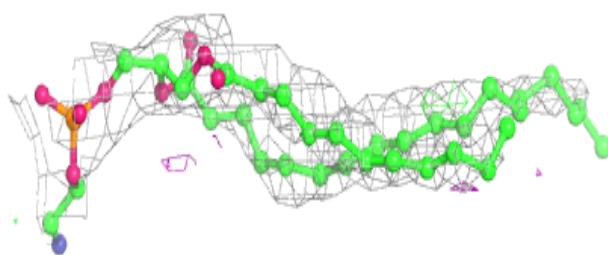
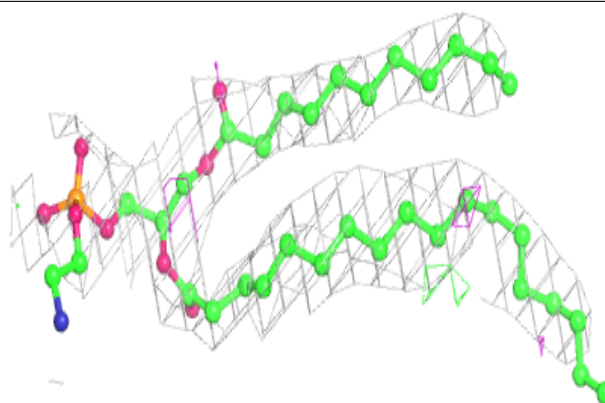
**Electron density around 3PE A 607:**

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

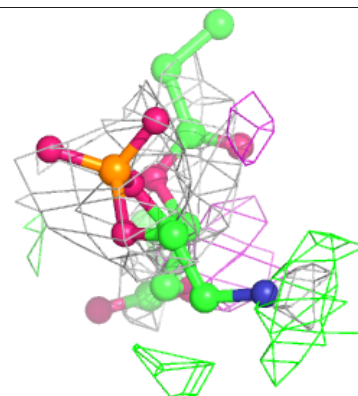
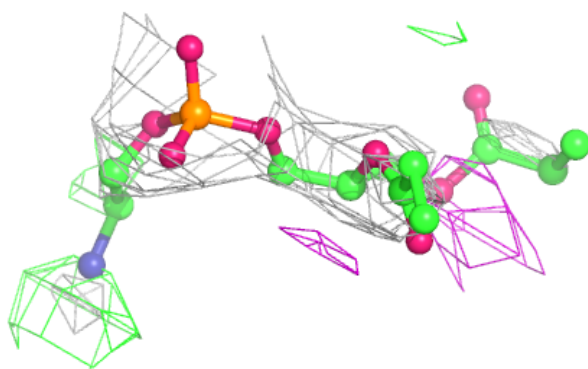
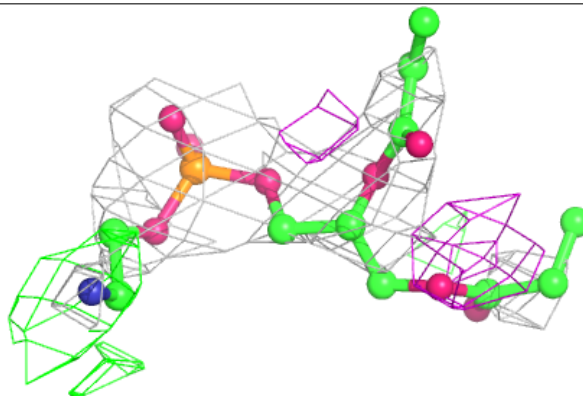


**Electron density around 3PE G 608:**

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

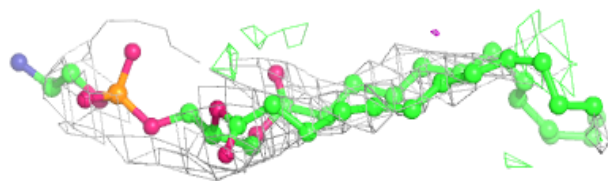
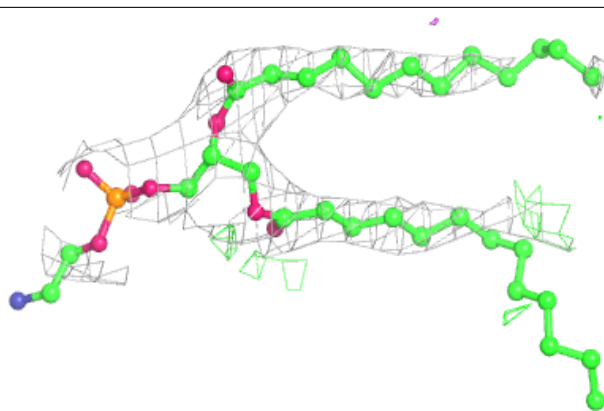
**Electron density around 3PE A 606:**

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

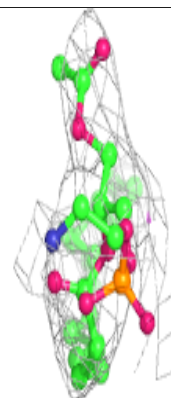
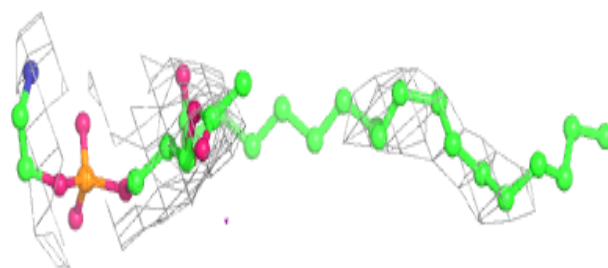
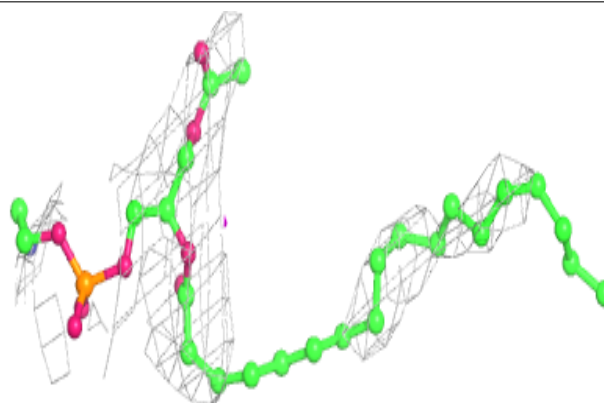


**Electron density around 3PE I 301:**

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

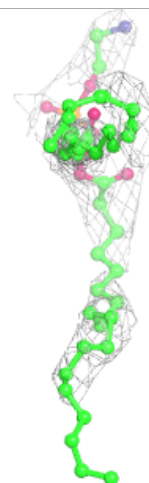
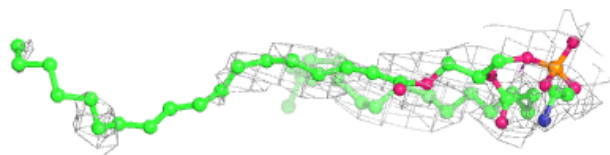
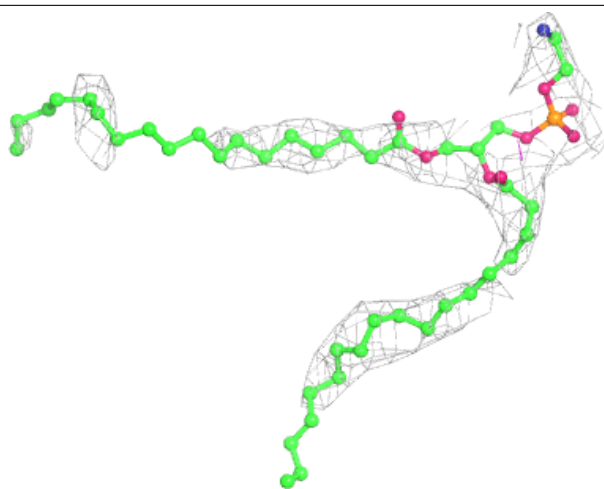
**Electron density around 3PE J 101:**

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



**Electron density around 3PE D 101:**

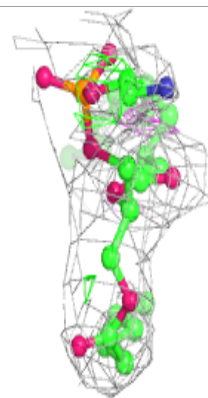
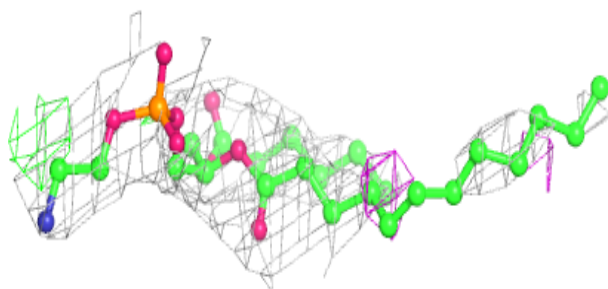
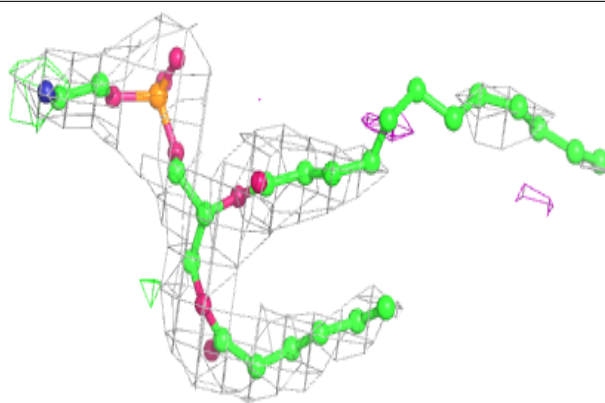
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



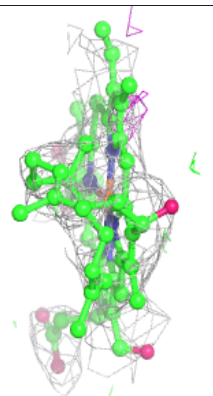
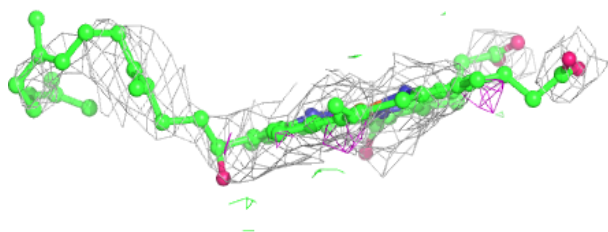
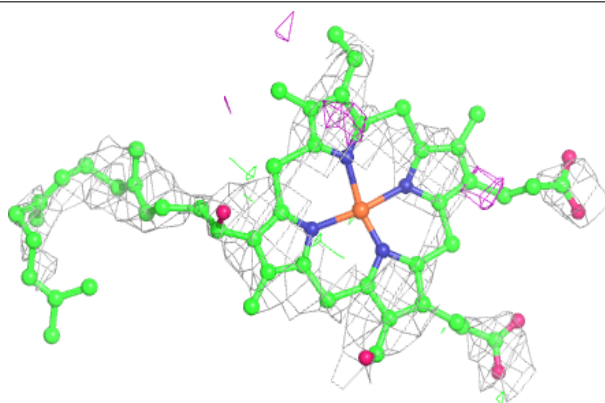


**Electron density around 3PE G 607:**

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

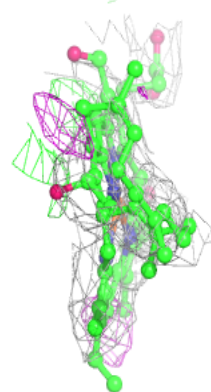
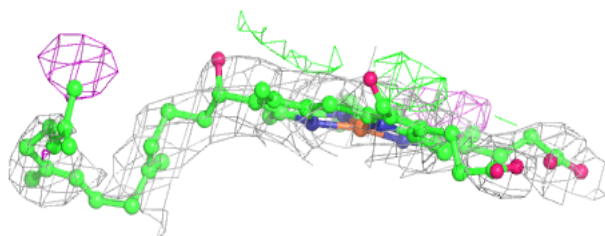
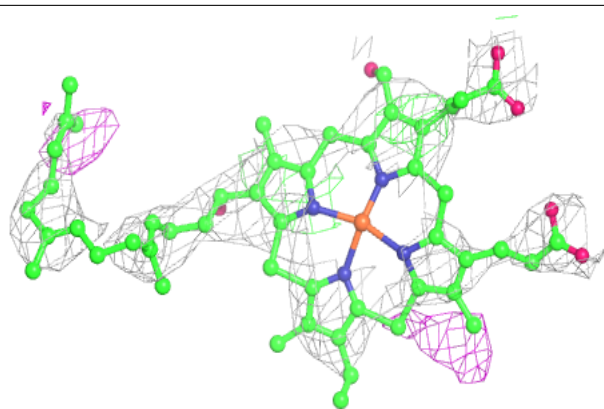
**Electron density around HEA G 603:**

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

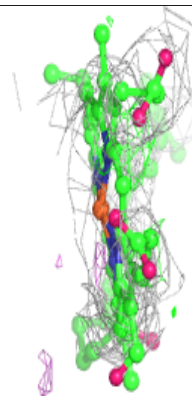
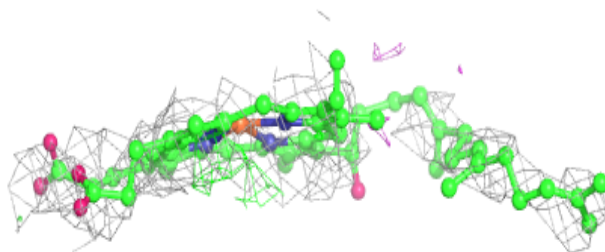
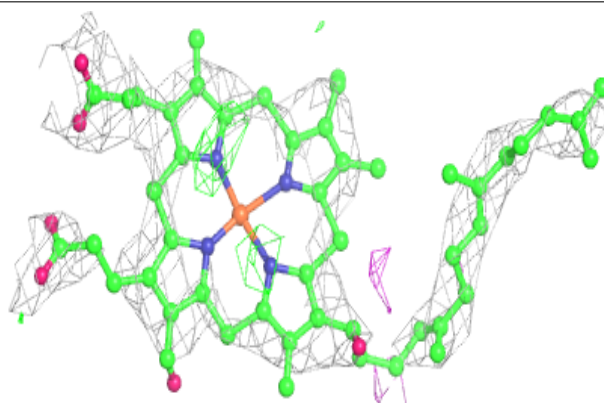


**Electron density around HEA A 602:**

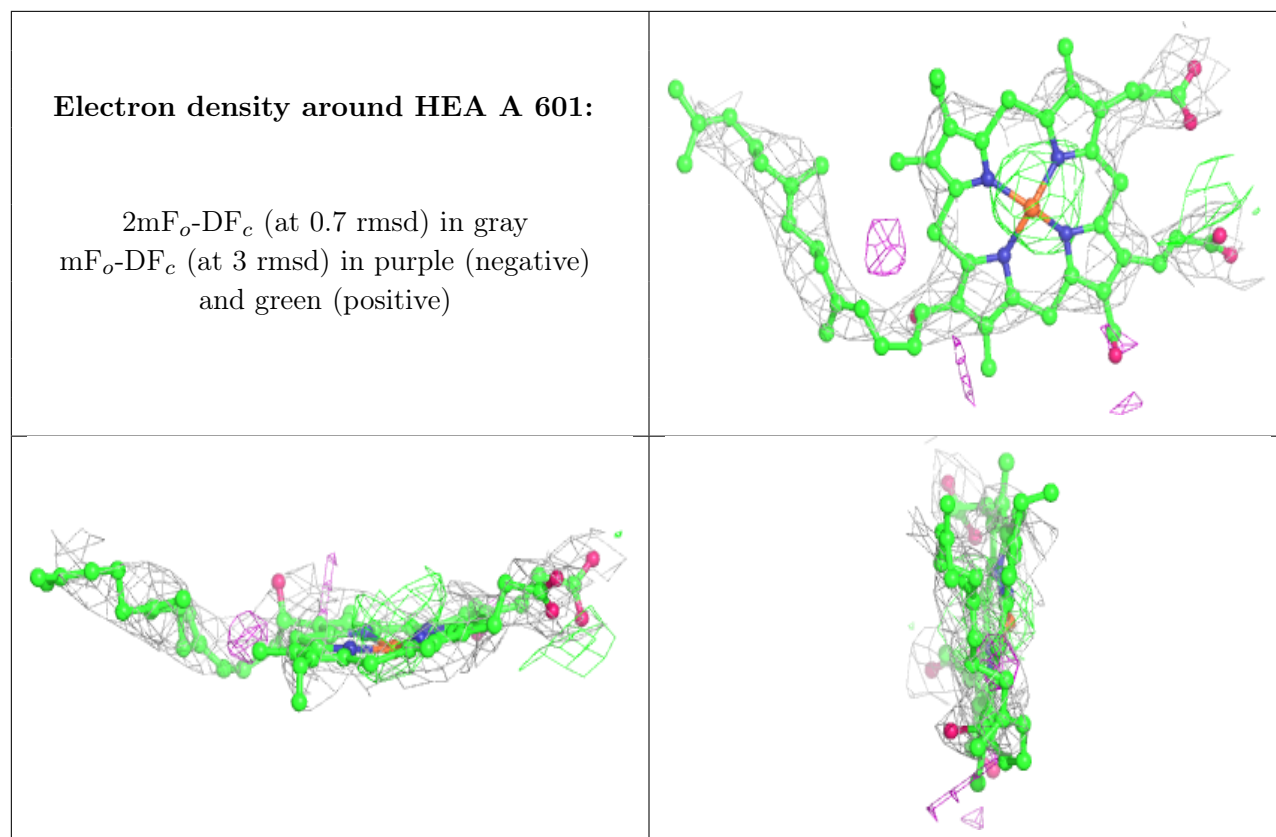
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around HEA G 602:**

$2mF_o-DF_c$  (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.