



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 17, 2023 – 11:12 AM EST

PDB ID : 4U8U  
Title : The Crystallographic structure of the giant hemoglobin from *Glossoscolex paulistus* at 3.2 Å resolution.  
Authors : Bachega, J.F.R.; Maluf, F.V.; Andi, B.; D'Muniz Pereira, H.; Carazzollea, M.F.; Orville, A.; Tabak, M.; Garratt, R.C.; Horjales, E.  
Deposited on : 2014-08-04  
Resolution : 3.20 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

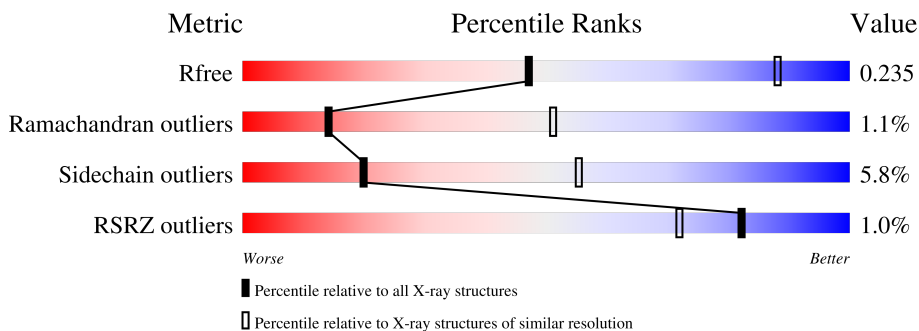
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	150	 90% 8% ..
1	E	150	 89% 9% .
1	I	150	 89% 9% .
1	P	150	 89% 10% .
1	T	150	 89% 9% ..
1	X	150	 91% 8% .

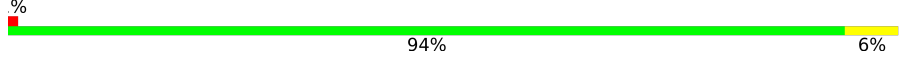
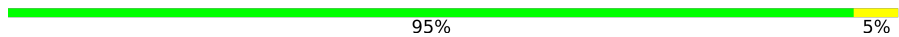
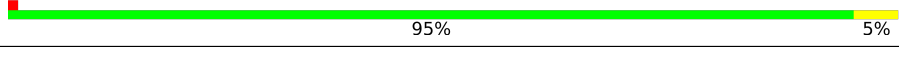
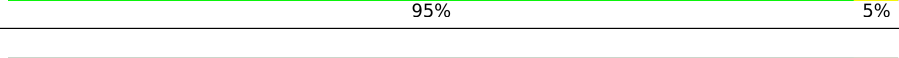
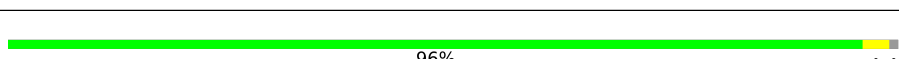
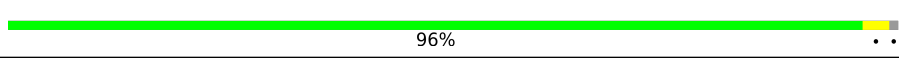
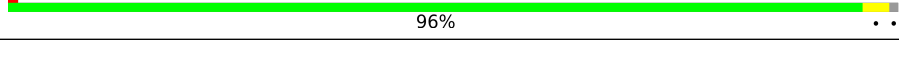

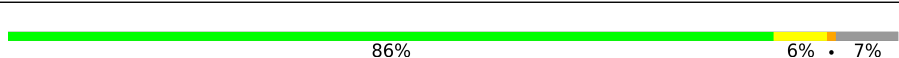
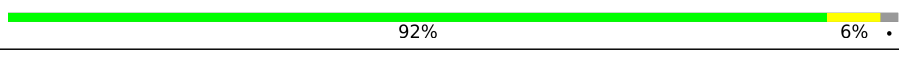
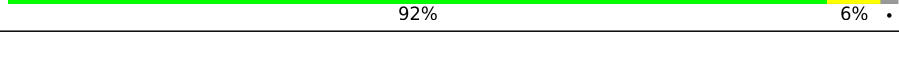
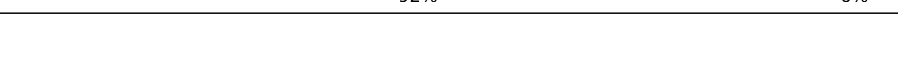


*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	e	150	3% 88% 10% ..
1	i	150	85% 12% ..
1	m	150	91% 8%
2	B	142	2% 94% 6%
2	F	142	% 94% 6%
2	J	142	2% 94% 6%
2	Q	142	% 94% 6%
2	U	142	% 94% 6%
2	Y	142	% 94% 6%
2	f	142	% 94% 6%
2	j	142	% 94% 6%
2	n	142	2% 94% 6%
3	C	151	% 97% .
3	G	151	3% 96% .
3	K	151	% 97% .
3	R	151	97% .
3	V	151	4% 97% .
3	Z	151	% 97% .
3	g	151	97% .
3	k	151	3% 97% .
3	o	151	2% 95% 5%
4	D	141	% 95% 5%
4	H	141	% 95% 5%
4	L	141	% 95% 5%
4	S	141	% 95% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	W	141	 % 94% 6%
4	a	141	 % 95% 5%
4	h	141	 % 95% 5%
4	l	141	 % 95% 5%
4	p	141	 % 95% 5%
5	M	224	 % 96% ..
5	b	224	 % 96% ..
5	q	224	 % 96% ..
6	N	236	 % 86% 6% • 7%
6	c	236	 % 86% 6% • 7%
6	r	236	 % 86% 6% • 7%
7	O	218	 % 92% 6% •
7	d	218	 % 92% 6% •
7	s	218	 % 92% 6% •

## 2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 59163 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 Globin a chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	148	1200	762	218	216	4	0	0	0
1	E	148	1200	762	218	216	4	0	0	0
1	I	148	1200	762	218	216	4	0	0	0
1	P	148	1200	762	218	216	4	0	0	0
1	T	148	1200	762	218	216	4	0	0	0
1	X	148	1200	762	218	216	4	0	0	0
1	e	148	1200	762	218	216	4	0	0	0
1	i	148	1200	762	218	216	4	0	0	0
1	m	148	1200	762	218	216	4	0	0	0

- Molecule 2 is a protein called Globin b Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	142	1148	727	215	203	3	0	0	0
2	F	142	1148	727	215	203	3	0	0	0
2	J	142	1148	727	215	203	3	0	0	0
2	Q	142	1148	727	215	203	3	0	0	0
2	U	142	1148	727	215	203	3	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	Y	142	Total	C	N	O	S	0	0	0
			1148	727	215	203	3			
2	f	142	Total	C	N	O	S	0	0	0
			1148	727	215	203	3			
2	j	142	Total	C	N	O	S	0	0	0
			1148	727	215	203	3			
2	n	142	Total	C	N	O	S	0	0	0
			1148	727	215	203	3			

- Molecule 3 is a protein called Globin c Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	151	Total	C	N	O	S	0	0	0
			1180	750	205	219	6			
3	G	151	Total	C	N	O	S	0	0	0
			1180	750	205	219	6			
3	K	151	Total	C	N	O	S	0	0	0
			1180	750	205	219	6			
3	R	151	Total	C	N	O	S	0	0	0
			1180	750	205	219	6			
3	V	151	Total	C	N	O	S	0	0	0
			1180	750	205	219	6			
3	Z	151	Total	C	N	O	S	0	0	0
			1180	750	205	219	6			
3	g	151	Total	C	N	O	S	0	0	0
			1180	750	205	219	6			
3	k	151	Total	C	N	O	S	0	0	0
			1180	750	205	219	6			
3	o	151	Total	C	N	O	S	0	0	0
			1189	755	208	220	6			

- Molecule 4 is a protein called Globin d Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	141	Total	C	N	O	S	0	0	0
			1140	734	199	204	3			
4	H	141	Total	C	N	O	S	0	0	0
			1140	734	199	204	3			
4	L	141	Total	C	N	O	S	0	0	0
			1140	734	199	204	3			
4	S	141	Total	C	N	O	S	0	0	0
			1140	734	199	204	3			

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	W	141	Total	C	N	O	S	0	0	0
			1140	734	199	204	3			
4	a	141	Total	C	N	O	S	0	0	0
			1140	734	199	204	3			
4	h	141	Total	C	N	O	S	0	0	0
			1140	734	199	204	3			
4	l	141	Total	C	N	O	S	0	0	0
			1140	734	199	204	3			
4	p	141	Total	C	N	O	S	0	0	0
			1140	734	199	204	3			

- Molecule 5 is a protein called Linker L1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	M	222	Total	C	N	O	S	0	0	0
			1754	1092	315	337	10			
5	b	222	Total	C	N	O	S	0	0	0
			1754	1092	315	337	10			
5	q	222	Total	C	N	O	S	0	0	0
			1754	1092	315	337	10			

- Molecule 6 is a protein called Linker L2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	N	219	Total	C	N	O	S	0	0	0
			1713	1060	326	317	10			
6	c	219	Total	C	N	O	S	0	0	0
			1713	1060	326	317	10			
6	r	219	Total	C	N	O	S	0	0	0
			1713	1060	326	317	10			

- Molecule 7 is a protein called Linker L3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	O	213	Total	C	N	O	S	0	0	0
			1686	1044	297	337	8			
7	d	213	Total	C	N	O	S	0	0	0
			1686	1044	297	337	8			
7	s	213	Total	C	N	O	S	0	0	0
			1686	1044	297	337	8			

- Molecule 8 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (for-





*Continued from previous page...*

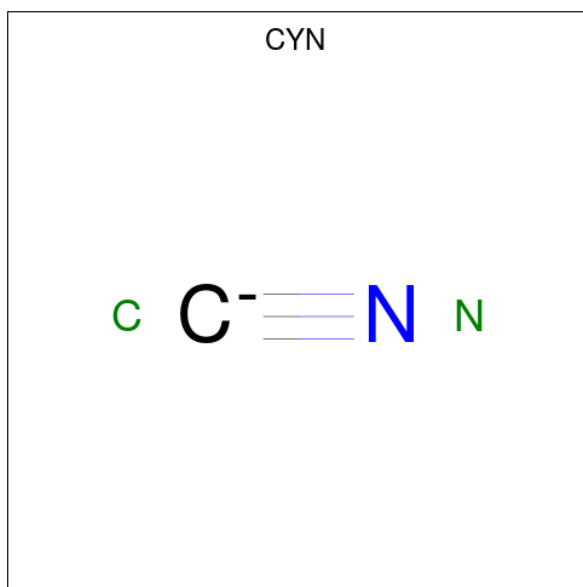
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
8	Q	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	R	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	S	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	T	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	U	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	V	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	W	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	X	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	Y	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	Z	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	a	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	e	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	f	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	g	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	h	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	i	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	j	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	k	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	l	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	m	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
8	n	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
8	o	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
8	p	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 9 is CYANIDE ION (three-letter code: CYN) (formula: CN).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total	C N	0	0
			2	1 1		
9	B	1	Total	C N	0	0
			2	1 1		
9	C	1	Total	C N	0	0
			2	1 1		
9	D	1	Total	C N	0	0
			2	1 1		
9	E	1	Total	C N	0	0
			2	1 1		
9	F	1	Total	C N	0	0
			2	1 1		
9	G	1	Total	C N	0	0
			2	1 1		
9	H	1	Total	C N	0	0
			2	1 1		
9	I	1	Total	C N	0	0
			2	1 1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	N		
9	J	1	2	1	1	0	0
9	K	1	2	1	1	0	0
9	L	1	2	1	1	0	0
9	P	1	2	1	1	0	0
9	Q	1	2	1	1	0	0
9	R	1	2	1	1	0	0
9	S	1	2	1	1	0	0
9	T	1	2	1	1	0	0
9	U	1	2	1	1	0	0
9	V	1	2	1	1	0	0
9	W	1	2	1	1	0	0
9	X	1	2	1	1	0	0
9	Y	1	2	1	1	0	0
9	Z	1	2	1	1	0	0
9	a	1	2	1	1	0	0
9	e	1	2	1	1	0	0
9	f	1	2	1	1	0	0
9	g	1	2	1	1	0	0
9	h	1	2	1	1	0	0
9	i	1	2	1	1	0	0
9	j	1	2	1	1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	k	1	Total C N 2 1 1	0	0
9	l	1	Total C N 2 1 1	0	0
9	m	1	Total C N 2 1 1	0	0
9	n	1	Total C N 2 1 1	0	0
9	o	1	Total C N 2 1 1	0	0
9	p	1	Total C N 2 1 1	0	0

- Molecule 10 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	M	1	Total Ca 1 1	0	0
10	N	1	Total Ca 1 1	0	0
10	O	1	Total Ca 1 1	0	0
10	b	1	Total Ca 1 1	0	0
10	c	1	Total Ca 1 1	0	0
10	d	1	Total Ca 1 1	0	0
10	q	1	Total Ca 1 1	0	0
10	r	1	Total Ca 1 1	0	0
10	s	1	Total Ca 1 1	0	0

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

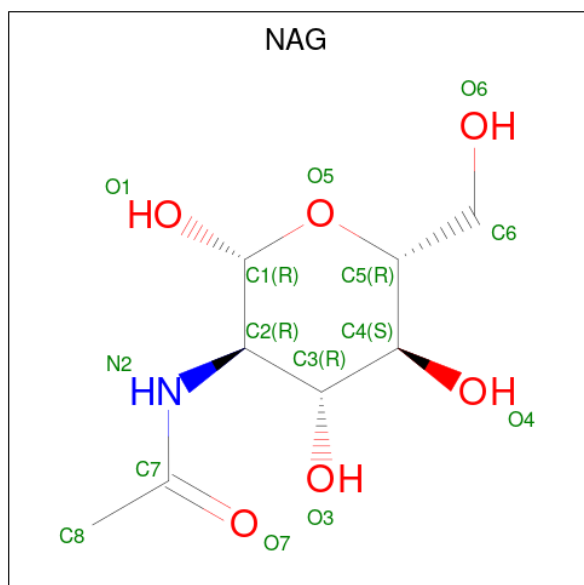
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	O	2	Total Zn 2 2	0	0
11	c	1	Total Zn 1 1	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	d	2	Total	Zn	0	0
			2	2		
11	s	1	Total	Zn	0	0
			1	1		

- Molecule 12 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
12	O	1	Total	C	N	O	0	0
			14	8	1	5		
12	d	1	Total	C	N	O	0	0
			14	8	1	5		
12	s	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 13 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	O	1	Total	O	0	0
			1	1		
13	c	1	Total	O	0	0
			1	1		
13	d	2	Total	O	0	0
			2	2		
13	r	1	Total	O	0	0
			1	1		

Continued on next page...

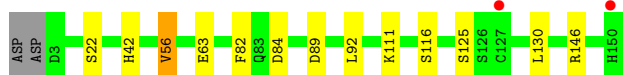
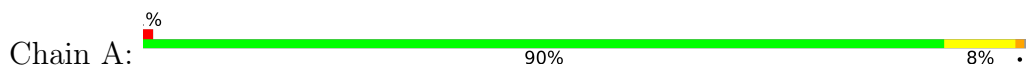
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
13	s	1	Total	O	0	0
			1	1		

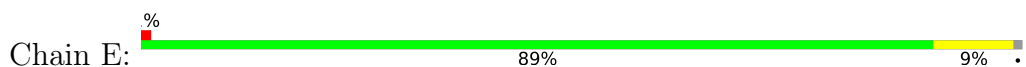
### 3 Residue-property plots [i](#)

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.

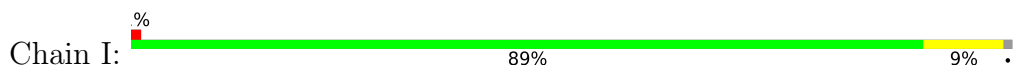
- Molecule 1: Globin a chain



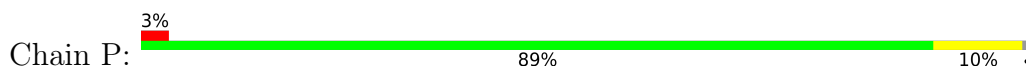
- Molecule 1: Globin a chain



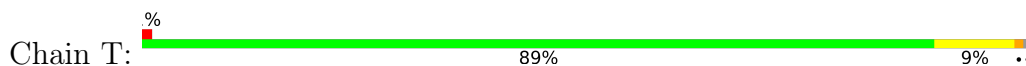
- Molecule 1: Globin a chain



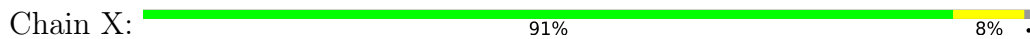
- Molecule 1: Globin a chain



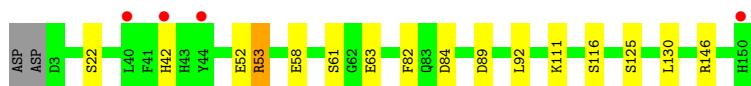
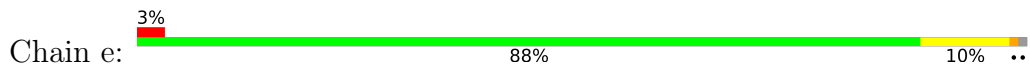
- Molecule 1: Globin a chain



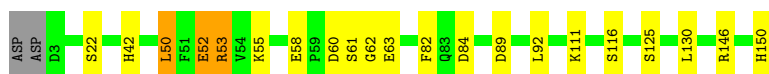
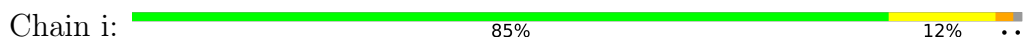
- Molecule 1: Globin a chain



● Molecule 1: Globin a chain



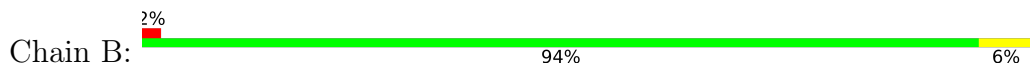
● Molecule 1: Globin a chain



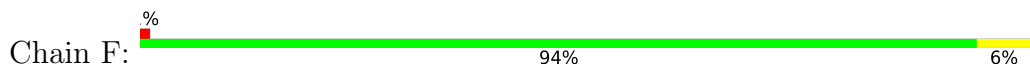
● Molecule 1: Globin a chain



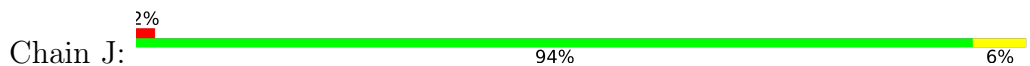
● Molecule 2: Globin b Chain



● Molecule 2: Globin b Chain



● Molecule 2: Globin b Chain

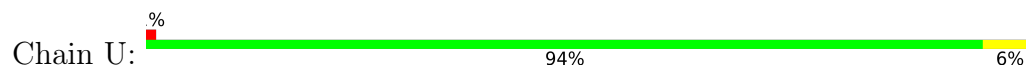


● Molecule 2: Globin b Chain

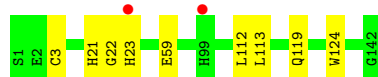




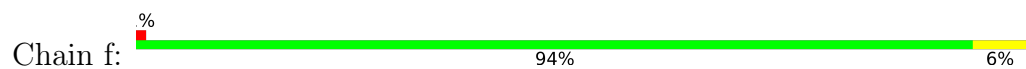
● Molecule 2: Globin b Chain



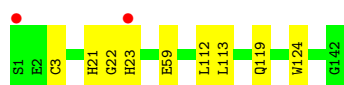
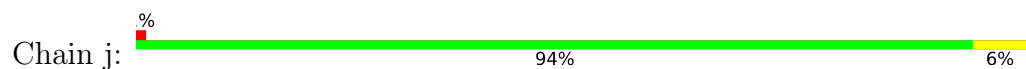
● Molecule 2: Globin b Chain



● Molecule 2: Globin b Chain



● Molecule 2: Globin b Chain



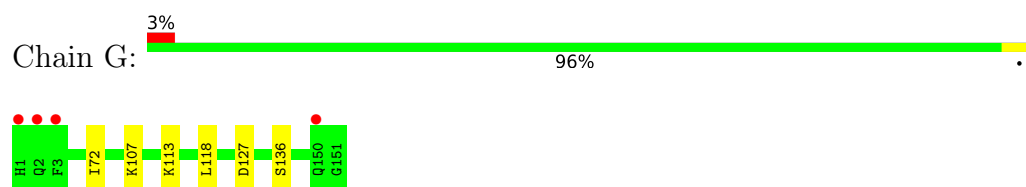
● Molecule 2: Globin b Chain



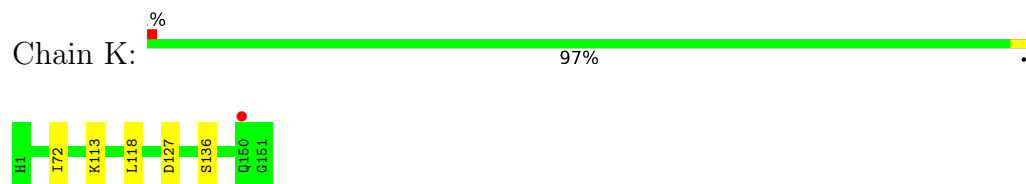
● Molecule 3: Globin c Chain



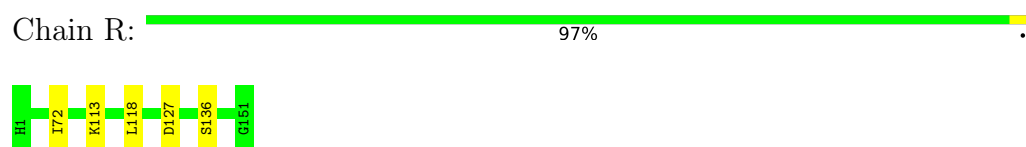
● Molecule 3: Globin c Chain



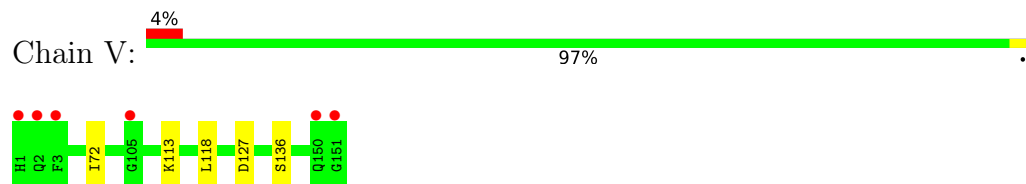
● Molecule 3: Globin c Chain



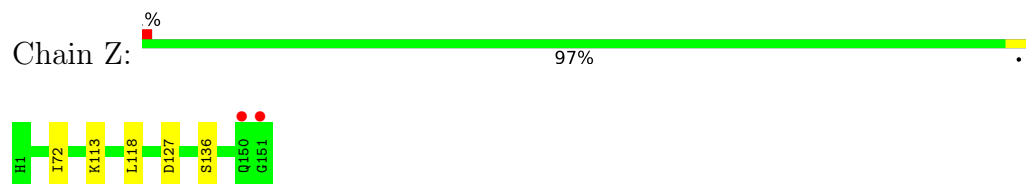
● Molecule 3: Globin c Chain



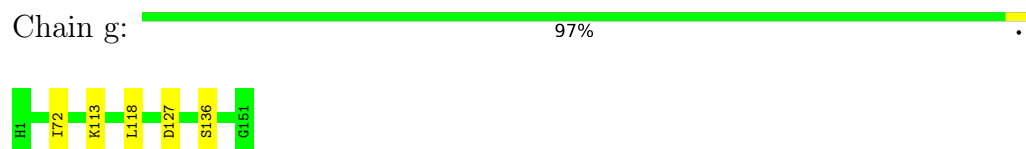
● Molecule 3: Globin c Chain



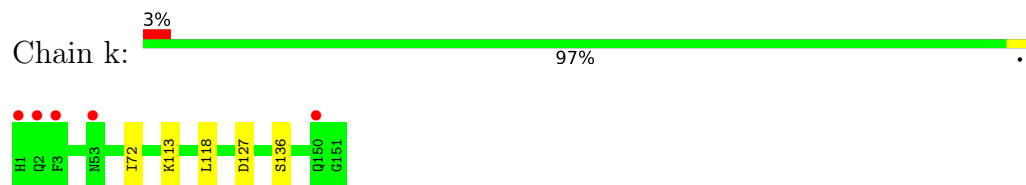
● Molecule 3: Globin c Chain



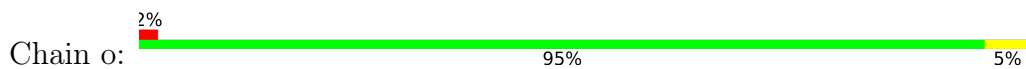
● Molecule 3: Globin c Chain



● Molecule 3: Globin c Chain



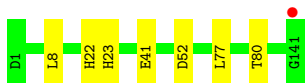
● Molecule 3: Globin c Chain



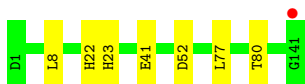
● Molecule 4: Globin d Chain



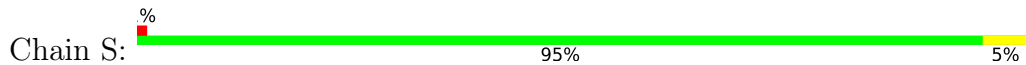
● Molecule 4: Globin d Chain



● Molecule 4: Globin d Chain



● Molecule 4: Globin d Chain



● Molecule 4: Globin d Chain



● Molecule 4: Globin d Chain



● Molecule 4: Globin d Chain



- Molecule 4: Globin d Chain



- Molecule 4: Globin d Chain



- Molecule 5: Linker L1



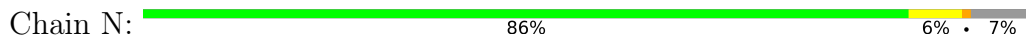
- Molecule 5: Linker L1




- Molecule 5: Linker L1



- Molecule 6: Linker L2




- Molecule 6: Linker L2

Chain c:  86% 6% • 7%



• Molecule 6: Linker L2

Chain r:  86% 6% • 7%



• Molecule 7: Linker L3

Chain O:  92% 6% • 7%



• Molecule 7: Linker L3

Chain d:  92% 6% • 7%



• Molecule 7: Linker L3

Chain s:  92% 6% • 7%



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	272.68Å 319.90Å 333.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.66 – 3.20 49.91 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.66-3.20) 94.1 (49.91-3.20)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.87 (at 3.19Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.215 , 0.235 0.216 , 0.235	Depositor DCC
$R_{free}$ test set	11885 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	54.2	Xtrriage
Anisotropy	0.312	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 26.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.007 for -h,-l,-k	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	59163	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.39% 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, CYN, HEM, ZN, 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	0.45	0/1231	0.52	0/1665
1	E	0.34	0/1231	0.50	0/1665
1	I	0.37	0/1231	0.95	3/1665 (0.2%)
1	P	0.33	0/1231	0.49	0/1665
1	T	0.49	0/1231	0.52	0/1665
1	X	0.59	0/1231	0.59	0/1665
1	e	0.53	0/1231	0.57	1/1665 (0.1%)
1	i	0.56	0/1231	0.63	2/1665 (0.1%)
1	m	0.37	0/1231	0.50	0/1665
2	B	0.32	0/1181	0.50	0/1600
2	F	0.32	0/1181	0.50	0/1600
2	J	0.31	0/1181	0.49	0/1600
2	Q	0.32	0/1181	0.50	0/1600
2	U	0.32	0/1181	0.51	0/1600
2	Y	0.32	0/1181	0.50	0/1600
2	f	0.32	0/1181	0.50	0/1600
2	j	0.31	0/1181	0.50	0/1600
2	n	0.32	0/1181	0.49	0/1600
3	C	0.29	0/1203	0.44	0/1622
3	G	0.29	0/1203	0.43	0/1622
3	K	0.28	0/1203	0.43	0/1622
3	R	0.29	0/1203	0.44	0/1622
3	V	0.28	0/1203	0.43	0/1622
3	Z	0.28	0/1203	0.43	0/1622
3	g	0.28	0/1203	0.43	0/1622
3	k	0.27	0/1203	0.43	0/1622
3	o	0.29	0/1213	0.44	0/1635
4	D	0.29	0/1172	0.46	0/1584
4	H	0.31	0/1172	0.46	0/1584
4	L	0.30	0/1172	0.46	0/1584
4	S	0.29	0/1172	0.46	0/1584
4	W	0.30	0/1172	0.46	0/1584

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
4	a	0.30	0/1172	0.46	0/1584
4	h	0.30	0/1172	0.46	0/1584
4	l	0.29	0/1172	0.46	0/1584
4	p	0.29	0/1172	0.46	0/1584
5	M	0.40	2/1793 (0.1%)	0.51	0/2421
5	b	0.35	0/1793	0.51	0/2421
5	q	0.34	0/1793	0.51	0/2421
6	N	0.32	0/1748	0.51	0/2357
6	c	0.32	0/1748	0.51	0/2357
6	r	0.33	0/1748	0.51	0/2357
7	O	0.32	0/1723	0.49	0/2328
7	d	0.33	0/1723	0.49	0/2328
7	s	0.32	0/1723	0.49	0/2328
All	All	0.34	2/58885 (0.0%)	0.50	6/79570 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
7	O	0	1
7	d	0	1
7	s	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	M	210	ASN	CG-ND2	-6.95	1.15	1.32
5	M	210	ASN	CG-OD1	-5.54	1.11	1.24

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	10	ARG	NE-CZ-NH1	-23.21	108.69	120.30
1	I	10	ARG	NE-CZ-NH2	21.13	130.87	120.30
1	I	10	ARG	CD-NE-CZ	10.18	137.85	123.60
1	i	50	LEU	CB-CG-CD2	-6.72	99.58	111.00
1	i	60	ASP	CB-CG-OD1	-6.12	112.80	118.30

There are no chirality outliers.



All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
7	O	94	HIS	Peptide
7	d	94	HIS	Peptide
7	s	94	HIS	Peptide

## 5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

## 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	146/150 (97%)	141 (97%)	3 (2%)	2 (1%)	11	46
1	E	146/150 (97%)	140 (96%)	4 (3%)	2 (1%)	11	46
1	I	146/150 (97%)	142 (97%)	3 (2%)	1 (1%)	22	61
1	P	146/150 (97%)	142 (97%)	3 (2%)	1 (1%)	22	61
1	T	146/150 (97%)	141 (97%)	3 (2%)	2 (1%)	11	46
1	X	146/150 (97%)	139 (95%)	6 (4%)	1 (1%)	22	61
1	e	146/150 (97%)	140 (96%)	4 (3%)	2 (1%)	11	46
1	i	146/150 (97%)	136 (93%)	7 (5%)	3 (2%)	7	37
1	m	146/150 (97%)	142 (97%)	3 (2%)	1 (1%)	22	61
2	B	140/142 (99%)	132 (94%)	5 (4%)	3 (2%)	7	37
2	F	140/142 (99%)	133 (95%)	4 (3%)	3 (2%)	7	37
2	J	140/142 (99%)	132 (94%)	5 (4%)	3 (2%)	7	37
2	Q	140/142 (99%)	132 (94%)	5 (4%)	3 (2%)	7	37
2	U	140/142 (99%)	133 (95%)	4 (3%)	3 (2%)	7	37
2	Y	140/142 (99%)	132 (94%)	5 (4%)	3 (2%)	7	37

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	f	140/142 (99%)	132 (94%)	5 (4%)	3 (2%)	7	37
2	j	140/142 (99%)	132 (94%)	5 (4%)	3 (2%)	7	37
2	n	140/142 (99%)	132 (94%)	5 (4%)	3 (2%)	7	37
3	C	149/151 (99%)	145 (97%)	4 (3%)	0	100	100
3	G	149/151 (99%)	145 (97%)	4 (3%)	0	100	100
3	K	149/151 (99%)	145 (97%)	4 (3%)	0	100	100
3	R	149/151 (99%)	145 (97%)	4 (3%)	0	100	100
3	V	149/151 (99%)	145 (97%)	4 (3%)	0	100	100
3	Z	149/151 (99%)	145 (97%)	4 (3%)	0	100	100
3	g	149/151 (99%)	145 (97%)	4 (3%)	0	100	100
3	k	149/151 (99%)	145 (97%)	4 (3%)	0	100	100
3	o	149/151 (99%)	145 (97%)	4 (3%)	0	100	100
4	D	139/141 (99%)	134 (96%)	5 (4%)	0	100	100
4	H	139/141 (99%)	134 (96%)	5 (4%)	0	100	100
4	L	139/141 (99%)	134 (96%)	5 (4%)	0	100	100
4	S	139/141 (99%)	134 (96%)	5 (4%)	0	100	100
4	W	139/141 (99%)	134 (96%)	5 (4%)	0	100	100
4	a	139/141 (99%)	134 (96%)	5 (4%)	0	100	100
4	h	139/141 (99%)	134 (96%)	5 (4%)	0	100	100
4	l	139/141 (99%)	134 (96%)	5 (4%)	0	100	100
4	p	139/141 (99%)	135 (97%)	4 (3%)	0	100	100
5	M	220/224 (98%)	208 (94%)	10 (4%)	2 (1%)	17	56
5	b	220/224 (98%)	208 (94%)	10 (4%)	2 (1%)	17	56
5	q	220/224 (98%)	207 (94%)	11 (5%)	2 (1%)	17	56
6	N	217/236 (92%)	198 (91%)	11 (5%)	8 (4%)	3	22
6	c	217/236 (92%)	199 (92%)	10 (5%)	8 (4%)	3	22
6	r	217/236 (92%)	198 (91%)	11 (5%)	8 (4%)	3	22
7	O	211/218 (97%)	199 (94%)	10 (5%)	2 (1%)	17	56
7	d	211/218 (97%)	200 (95%)	9 (4%)	2 (1%)	17	56
7	s	211/218 (97%)	199 (94%)	10 (5%)	2 (1%)	17	56
All	All	7110/7290 (98%)	6781 (95%)	251 (4%)	78 (1%)	14	51

5 of 78 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	N	19	PRO
6	N	76	CYS
6	N	233	GLN
6	N	235	HIS
6	c	19	PRO

### 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	129/131 (98%)	117 (91%)	12 (9%)	9 33
1	E	129/131 (98%)	117 (91%)	12 (9%)	9 33
1	I	129/131 (98%)	117 (91%)	12 (9%)	9 33
1	P	129/131 (98%)	116 (90%)	13 (10%)	7 29
1	T	129/131 (98%)	116 (90%)	13 (10%)	7 29
1	X	129/131 (98%)	118 (92%)	11 (8%)	10 38
1	e	129/131 (98%)	115 (89%)	14 (11%)	6 26
1	i	129/131 (98%)	111 (86%)	18 (14%)	3 16
1	m	129/131 (98%)	118 (92%)	11 (8%)	10 38
2	B	117/117 (100%)	111 (95%)	6 (5%)	24 60
2	F	117/117 (100%)	111 (95%)	6 (5%)	24 60
2	J	117/117 (100%)	111 (95%)	6 (5%)	24 60
2	Q	117/117 (100%)	111 (95%)	6 (5%)	24 60
2	U	117/117 (100%)	111 (95%)	6 (5%)	24 60
2	Y	117/117 (100%)	111 (95%)	6 (5%)	24 60
2	f	117/117 (100%)	111 (95%)	6 (5%)	24 60
2	j	117/117 (100%)	111 (95%)	6 (5%)	24 60
2	n	117/117 (100%)	111 (95%)	6 (5%)	24 60
3	C	128/130 (98%)	123 (96%)	5 (4%)	32 67

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	G	128/130 (98%)	123 (96%)	5 (4%)	32	67
3	K	128/130 (98%)	123 (96%)	5 (4%)	32	67
3	R	128/130 (98%)	123 (96%)	5 (4%)	32	67
3	V	128/130 (98%)	123 (96%)	5 (4%)	32	67
3	Z	128/130 (98%)	123 (96%)	5 (4%)	32	67
3	g	128/130 (98%)	123 (96%)	5 (4%)	32	67
3	k	128/130 (98%)	123 (96%)	5 (4%)	32	67
3	o	130/130 (100%)	122 (94%)	8 (6%)	18	53
4	D	120/120 (100%)	113 (94%)	7 (6%)	20	55
4	H	120/120 (100%)	113 (94%)	7 (6%)	20	55
4	L	120/120 (100%)	113 (94%)	7 (6%)	20	55
4	S	120/120 (100%)	113 (94%)	7 (6%)	20	55
4	W	120/120 (100%)	113 (94%)	7 (6%)	20	55
4	a	120/120 (100%)	113 (94%)	7 (6%)	20	55
4	h	120/120 (100%)	113 (94%)	7 (6%)	20	55
4	l	120/120 (100%)	113 (94%)	7 (6%)	20	55
4	p	120/120 (100%)	113 (94%)	7 (6%)	20	55
5	M	185/194 (95%)	181 (98%)	4 (2%)	52	79
5	b	185/194 (95%)	181 (98%)	4 (2%)	52	79
5	q	185/194 (95%)	181 (98%)	4 (2%)	52	79
6	N	181/205 (88%)	171 (94%)	10 (6%)	21	57
6	c	181/205 (88%)	171 (94%)	10 (6%)	21	57
6	r	181/205 (88%)	171 (94%)	10 (6%)	21	57
7	O	187/193 (97%)	177 (95%)	10 (5%)	22	58
7	d	187/193 (97%)	177 (95%)	10 (5%)	22	58
7	s	187/193 (97%)	177 (95%)	10 (5%)	22	58
All	All	6107/6258 (98%)	5754 (94%)	353 (6%)	20	55

5 of 353 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	e	58	GLU
3	k	113	LYS

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	e	111	LYS
4	h	80	THR
1	m	82	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
5	b	195	HIS
6	c	231	HIS
5	q	195	HIS
1	i	150	HIS
1	m	150	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 [i](#)

Of 90 ligands modelled in this entry, 15 are monoatomic - leaving 75 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	$\# Z  > 2$
9	CYN	Z	202	-	0,1,1	-	-	-		

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	HEM	E	201	1	41,50,50	2.22	9 (21%)	45,82,82	2.04	11 (24%)
8	HEM	R	201	3	41,50,50	2.18	8 (19%)	45,82,82	1.95	12 (26%)
8	HEM	D	201	4	41,50,50	2.22	8 (19%)	45,82,82	1.91	12 (26%)
9	CYN	m	202	8	0,1,1	-	-	-		
9	CYN	a	202	8	0,1,1	-	-	-		
9	CYN	B	202	8	0,1,1	-	-	-		
9	CYN	L	202	-	0,1,1	-	-	-		
12	NAG	s	303	7	14,14,15	0.48	0	17,19,21	1.35	3 (17%)
9	CYN	f	202	8	0,1,1	-	-	-		
9	CYN	n	202	8	0,1,1	-	-	-		
8	HEM	X	201	1	41,50,50	2.26	9 (21%)	45,82,82	2.18	13 (28%)
9	CYN	F	202	8	0,1,1	-	-	-		
9	CYN	A	202	-	0,1,1	-	-	-		
9	CYN	o	202	8	0,1,1	-	-	-		
9	CYN	D	202	-	0,1,1	-	-	-		
8	HEM	l	201	4	41,50,50	2.24	8 (19%)	45,82,82	1.91	12 (26%)
9	CYN	p	202	8	0,1,1	-	-	-		
12	NAG	d	304	7	14,14,15	0.43	0	17,19,21	1.19	2 (11%)
9	CYN	U	202	8	0,1,1	-	-	-		
8	HEM	j	201	2,9	41,50,50	2.22	9 (21%)	45,82,82	1.95	13 (28%)
9	CYN	S	202	-	0,1,1	-	-	-		
9	CYN	W	202	8	0,1,1	-	-	-		
9	CYN	X	202	-	0,1,1	-	-	-		
8	HEM	e	201	9,1	41,50,50	2.22	7 (17%)	45,82,82	2.17	18 (40%)
8	HEM	A	201	1	41,50,50	2.21	7 (17%)	45,82,82	2.09	15 (33%)
8	HEM	n	201	2,9	41,50,50	2.24	9 (21%)	45,82,82	1.97	13 (28%)
9	CYN	Q	202	-	0,1,1	-	-	-		
9	CYN	j	202	8	0,1,1	-	-	-		
9	CYN	k	202	8	0,1,1	-	-	-		
9	CYN	l	202	-	0,1,1	-	-	-		
8	HEM	V	201	3,9	41,50,50	2.20	8 (19%)	45,82,82	1.96	12 (26%)
9	CYN	h	202	-	0,1,1	-	-	-		
9	CYN	I	202	-	0,1,1	-	-	-		
9	CYN	H	202	-	0,1,1	-	-	-		
9	CYN	g	202	-	0,1,1	-	-	-		
9	CYN	i	202	-	0,1,1	-	-	-		
9	CYN	G	202	8	0,1,1	-	-	-		
8	HEM	J	201	2,9	41,50,50	2.25	9 (21%)	45,82,82	1.99	14 (31%)
8	HEM	i	201	1	41,50,50	2.26	10 (24%)	45,82,82	2.00	13 (28%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	HEM	g	201	3	41,50,50	2.19	8 (19%)	45,82,82	1.88	11 (24%)
8	HEM	G	201	3,9	41,50,50	2.19	7 (17%)	45,82,82	2.02	13 (28%)
9	CYN	E	202	-	0,1,1	-	-	-	-	-
8	HEM	F	201	2,9	41,50,50	2.21	7 (17%)	45,82,82	1.95	13 (28%)
8	HEM	T	201	1	41,50,50	2.26	10 (24%)	45,82,82	2.00	13 (28%)
8	HEM	P	201	1	41,50,50	2.27	9 (21%)	45,82,82	2.11	17 (37%)
9	CYN	R	202	-	0,1,1	-	-	-	-	-
8	HEM	K	201	3	41,50,50	2.20	8 (19%)	45,82,82	2.01	12 (26%)
8	HEM	Z	201	3	41,50,50	2.20	8 (19%)	45,82,82	1.89	12 (26%)
8	HEM	m	201	9,1	41,50,50	2.26	10 (24%)	45,82,82	2.22	16 (35%)
8	HEM	o	201	3,9	41,50,50	2.19	7 (17%)	45,82,82	1.96	12 (26%)
8	HEM	B	201	2,9	41,50,50	2.21	8 (19%)	45,82,82	1.88	12 (26%)
8	HEM	L	201	4	41,50,50	2.22	10 (24%)	45,82,82	1.95	13 (28%)
9	CYN	K	202	-	0,1,1	-	-	-	-	-
8	HEM	f	201	2,9	41,50,50	2.19	9 (21%)	45,82,82	1.89	12 (26%)
9	CYN	P	202	-	0,1,1	-	-	-	-	-
8	HEM	Y	201	2,9	41,50,50	2.22	8 (19%)	45,82,82	1.90	11 (24%)
8	HEM	C	201	3	41,50,50	2.16	8 (19%)	45,82,82	1.82	12 (26%)
8	HEM	p	201	4,9	41,50,50	2.25	10 (24%)	45,82,82	2.09	12 (26%)
9	CYN	Y	202	8	0,1,1	-	-	-	-	-
8	HEM	S	201	4	41,50,50	2.23	8 (19%)	45,82,82	1.93	12 (26%)
8	HEM	U	201	2,9	41,50,50	2.22	7 (17%)	45,82,82	1.96	12 (26%)
9	CYN	V	202	8	0,1,1	-	-	-	-	-
9	CYN	e	202	8	0,1,1	-	-	-	-	-
8	HEM	Q	201	2	41,50,50	2.21	8 (19%)	45,82,82	1.90	12 (26%)
8	HEM	H	201	4	41,50,50	2.21	8 (19%)	45,82,82	1.85	12 (26%)
8	HEM	W	201	4,9	41,50,50	2.24	7 (17%)	45,82,82	2.00	14 (31%)
8	HEM	k	201	3,9	41,50,50	2.25	8 (19%)	45,82,82	2.03	13 (28%)
9	CYN	C	202	-	0,1,1	-	-	-	-	-
12	NAG	O	304	7	14,14,15	0.40	0	17,19,21	1.77	3 (17%)
8	HEM	h	201	4	41,50,50	2.23	8 (19%)	45,82,82	1.85	10 (22%)
8	HEM	I	201	1	41,50,50	2.22	9 (21%)	45,82,82	2.05	16 (35%)
9	CYN	T	202	-	0,1,1	-	-	-	-	-
9	CYN	J	202	8	0,1,1	-	-	-	-	-
8	HEM	a	201	4,9	41,50,50	2.22	8 (19%)	45,82,82	2.00	11 (24%)

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
8	HEM	E	201	1	-	4/12/54/54	-
8	HEM	R	201	3	-	4/12/54/54	-
8	HEM	D	201	4	-	4/12/54/54	-
12	NAG	s	303	7	-	0/6/23/26	0/1/1/1
8	HEM	X	201	1	-	5/12/54/54	-
8	HEM	l	201	4	-	7/12/54/54	-
12	NAG	d	304	7	-	0/6/23/26	0/1/1/1
8	HEM	j	201	2,9	-	6/12/54/54	-
8	HEM	e	201	9,1	-	6/12/54/54	-
8	HEM	A	201	1	-	4/12/54/54	-
8	HEM	n	201	2,9	-	4/12/54/54	-
8	HEM	V	201	3,9	-	6/12/54/54	-
8	HEM	J	201	2,9	-	4/12/54/54	-
8	HEM	i	201	1	-	4/12/54/54	-
8	HEM	g	201	3	-	4/12/54/54	-
8	HEM	G	201	3,9	-	4/12/54/54	-
8	HEM	F	201	2,9	-	6/12/54/54	-
8	HEM	T	201	1	-	4/12/54/54	-
8	HEM	P	201	1	-	5/12/54/54	-
8	HEM	K	201	3	-	4/12/54/54	-
8	HEM	Z	201	3	-	3/12/54/54	-
8	HEM	m	201	9,1	-	4/12/54/54	-
8	HEM	o	201	3,9	-	4/12/54/54	-
8	HEM	B	201	2,9	-	6/12/54/54	-
8	HEM	L	201	4	-	8/12/54/54	-
8	HEM	f	201	2,9	-	4/12/54/54	-
8	HEM	Y	201	2,9	-	8/12/54/54	-
8	HEM	C	201	3	-	6/12/54/54	-
8	HEM	p	201	4,9	-	6/12/54/54	-
8	HEM	S	201	4	-	6/12/54/54	-
8	HEM	U	201	2,9	-	4/12/54/54	-
8	HEM	Q	201	2	-	6/12/54/54	-
8	HEM	H	201	4	-	4/12/54/54	-

*Continued on next page...*



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	HEM	W	201	4,9	-	6/12/54/54	-
8	HEM	k	201	3,9	-	4/12/54/54	-
12	NAG	O	304	7	-	0/6/23/26	0/1/1/1
8	HEM	h	201	4	-	6/12/54/54	-
8	HEM	I	201	1	-	2/12/54/54	-
8	HEM	a	201	4,9	-	6/12/54/54	-

The worst 5 of 299 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	l	201	HEM	C3D-C2D	8.41	1.54	1.36
8	J	201	HEM	C3D-C2D	8.38	1.54	1.36
8	p	201	HEM	C3D-C2D	8.37	1.54	1.36
8	W	201	HEM	C3D-C2D	8.28	1.54	1.36
8	D	201	HEM	C3D-C2D	8.28	1.54	1.36

The worst 5 of 469 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	e	201	HEM	C4D-ND-C1D	7.63	112.95	105.07
8	P	201	HEM	C4D-ND-C1D	7.60	112.92	105.07
8	E	201	HEM	C4D-ND-C1D	7.44	112.76	105.07
8	i	201	HEM	C4D-ND-C1D	7.39	112.70	105.07
8	A	201	HEM	C4D-ND-C1D	7.37	112.69	105.07

There are no chirality outliers.

5 of 178 torsion outliers are listed below:

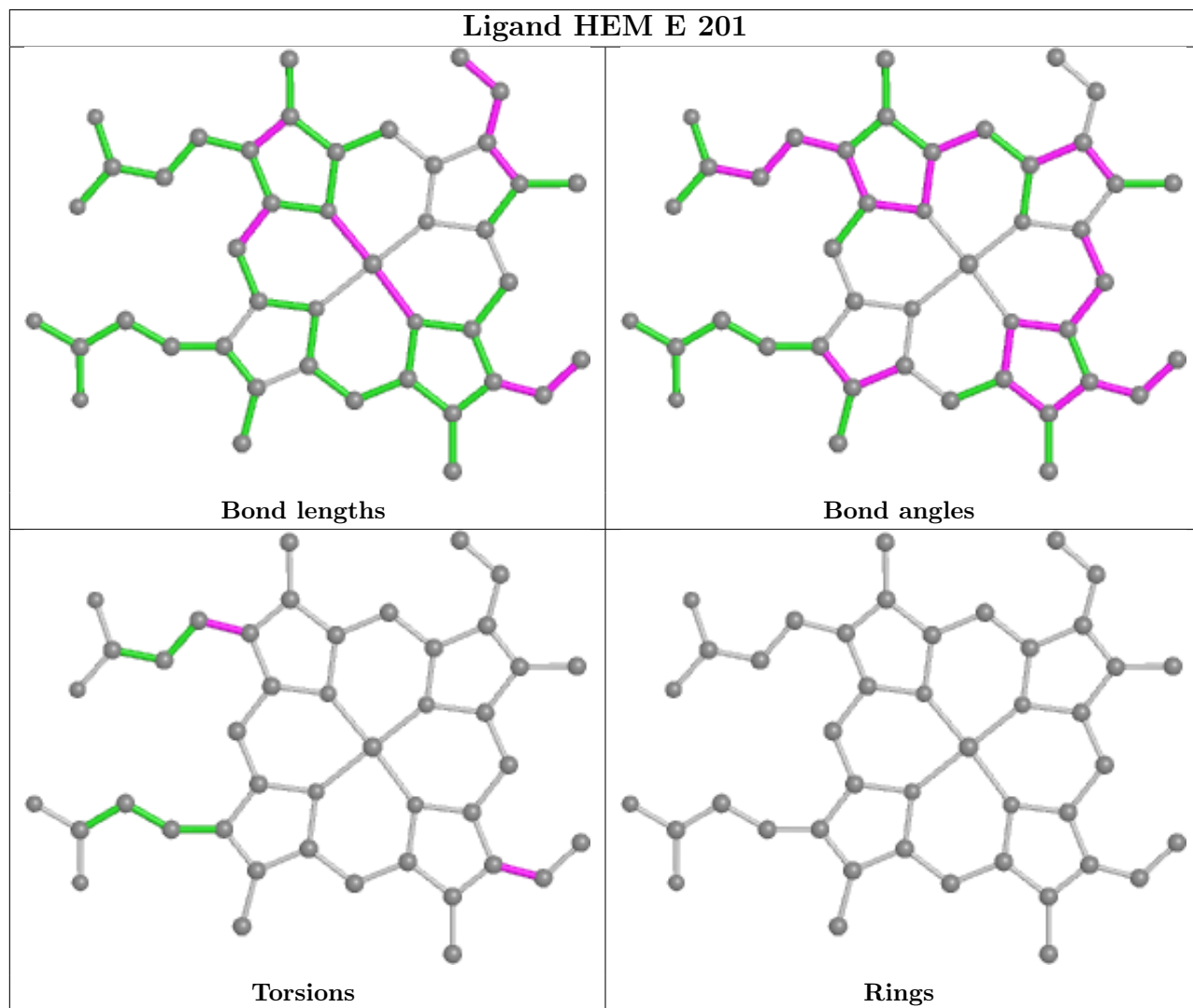
Mol	Chain	Res	Type	Atoms
8	A	201	HEM	C2B-C3B-CAB-CBB
8	A	201	HEM	C4B-C3B-CAB-CBB
8	B	201	HEM	C2B-C3B-CAB-CBB
8	C	201	HEM	C2B-C3B-CAB-CBB
8	C	201	HEM	C4B-C3B-CAB-CBB

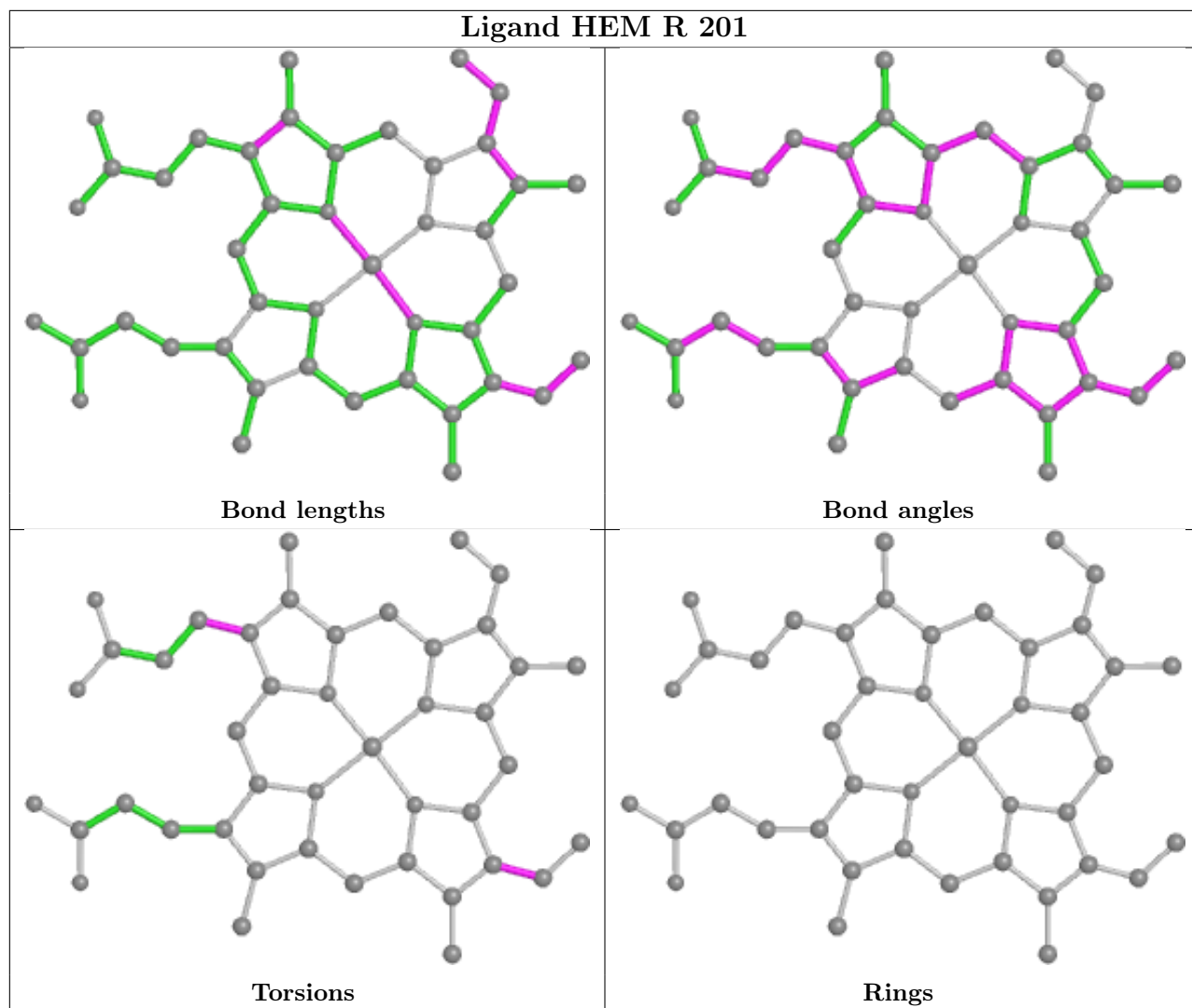
There are no ring outliers.

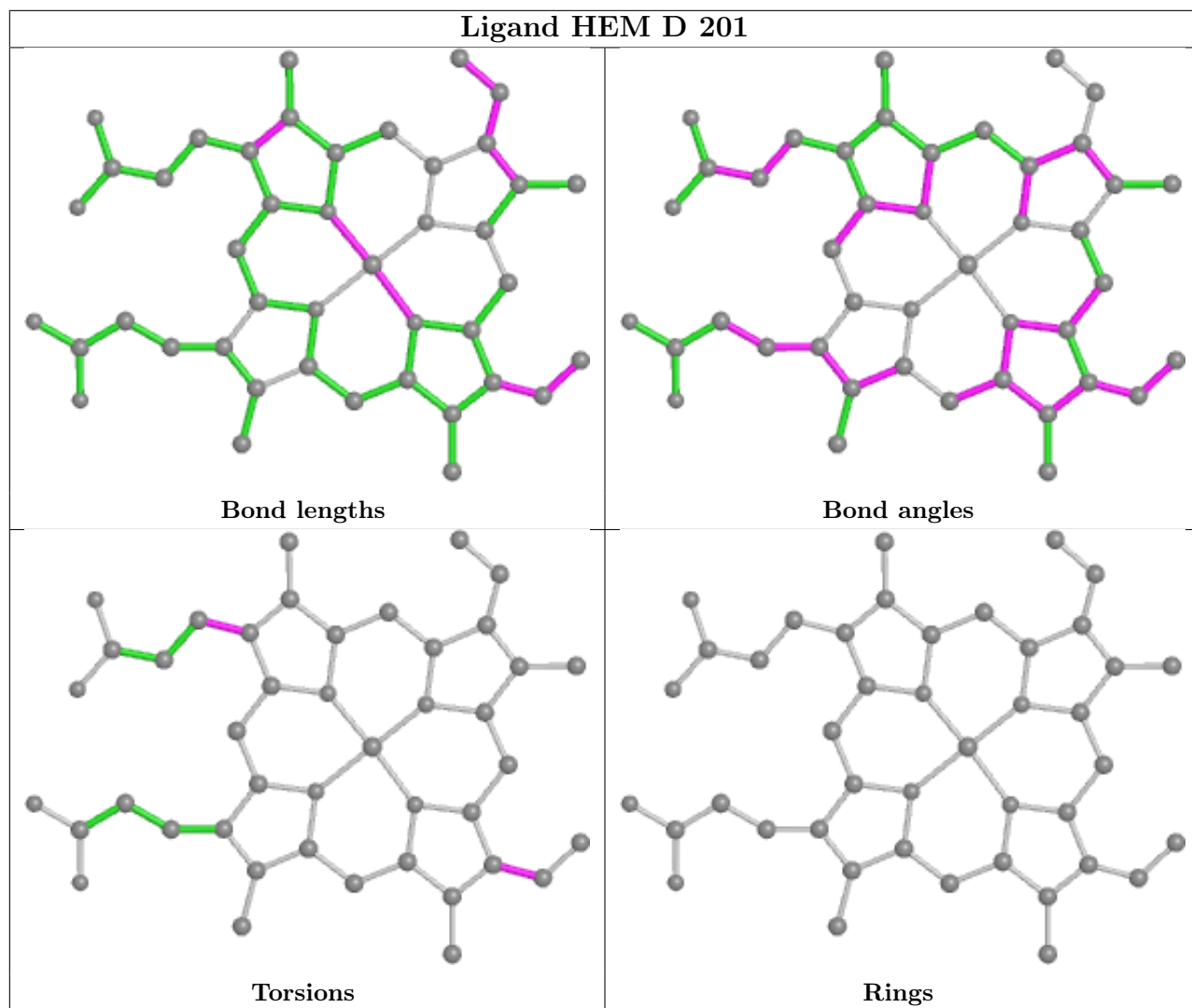
No monomer is involved in short contacts.

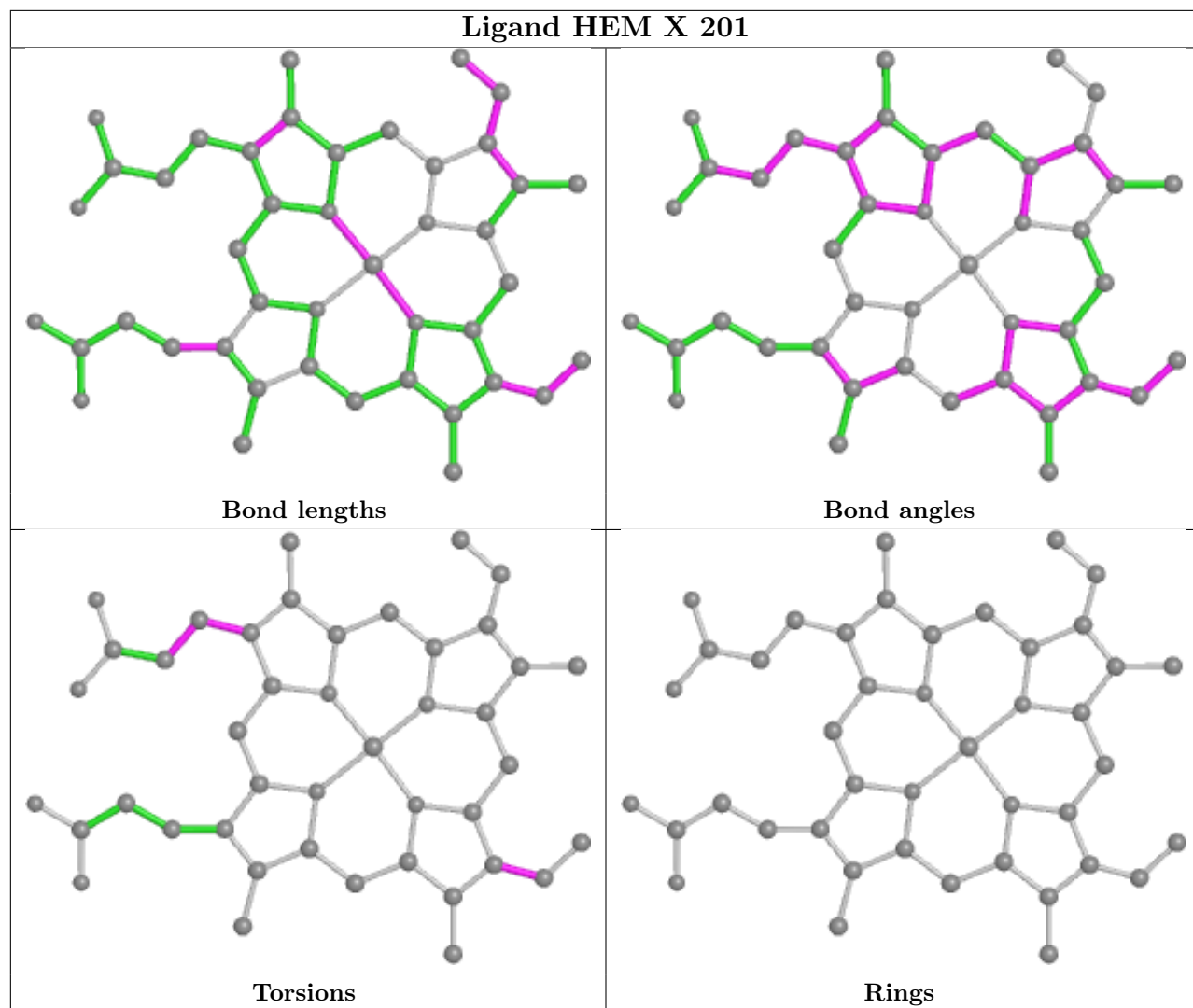
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

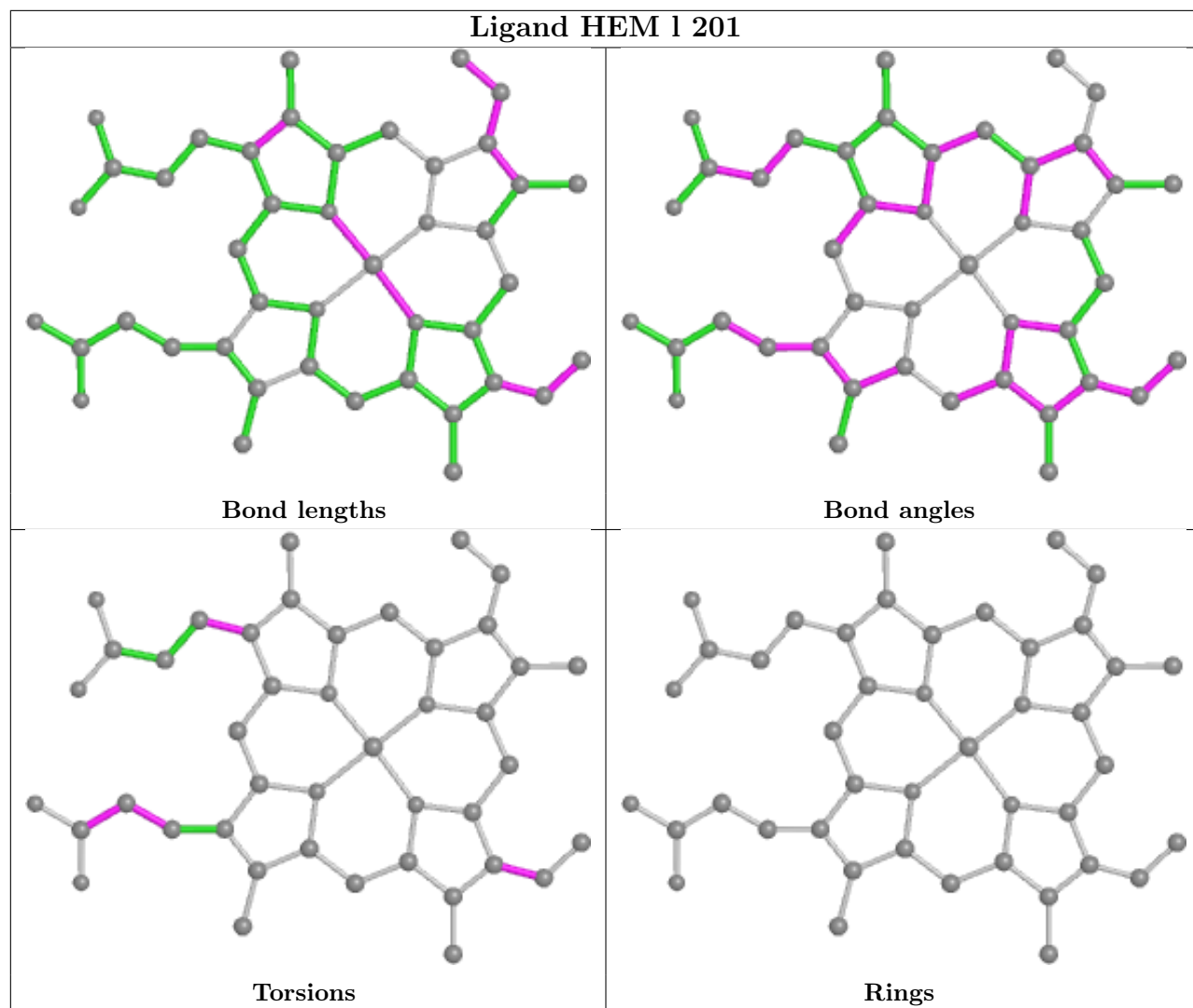
also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted 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.

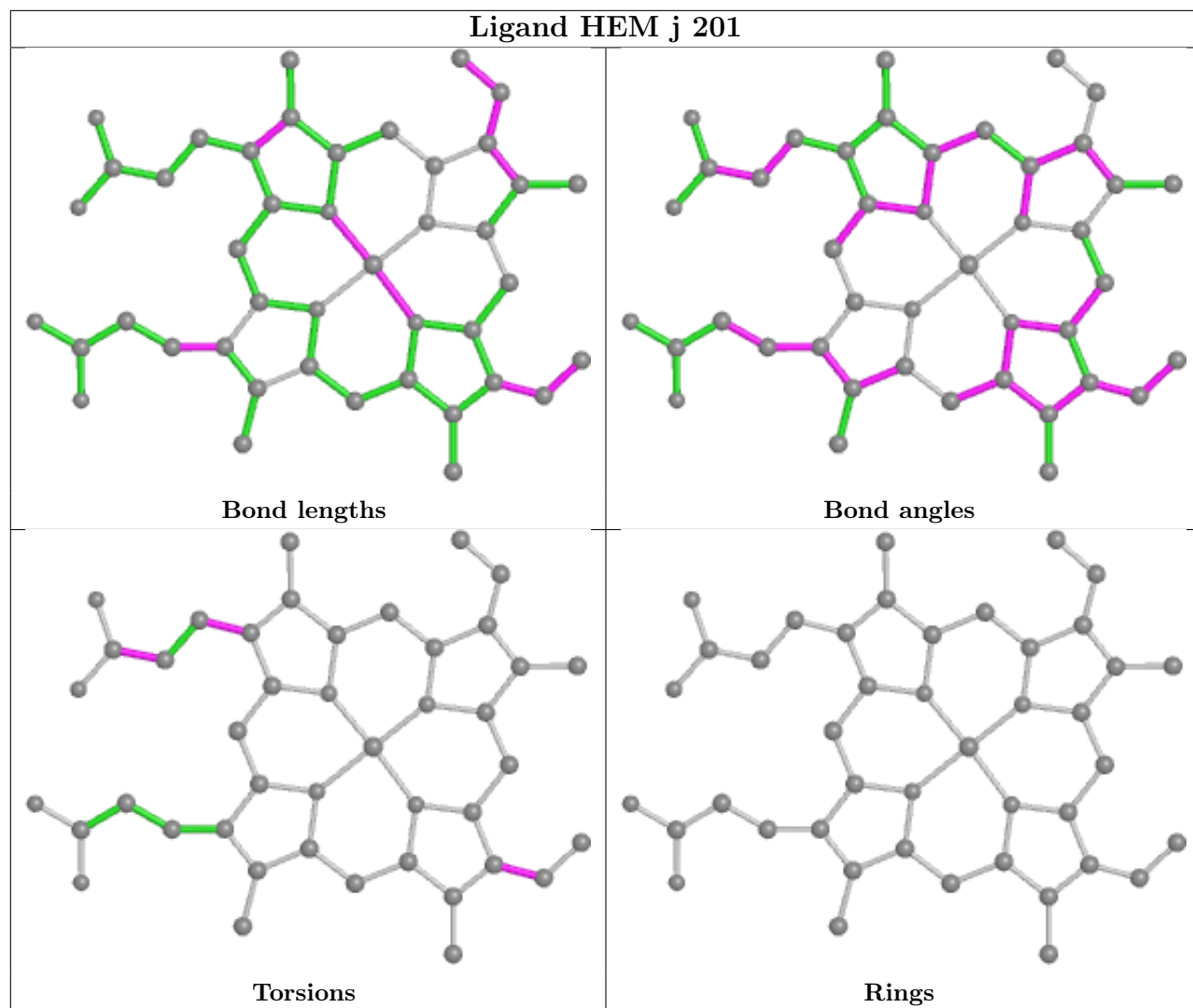


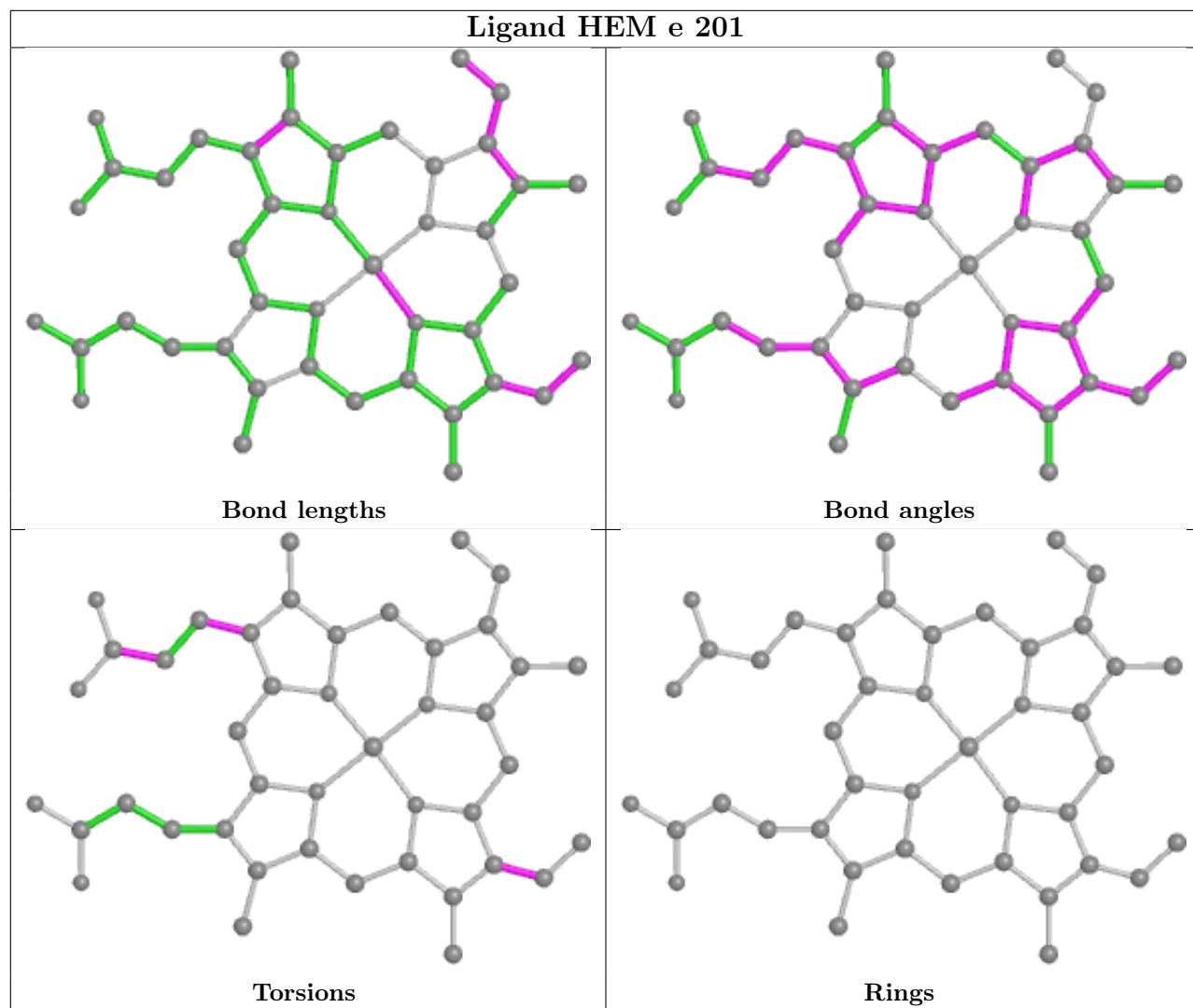




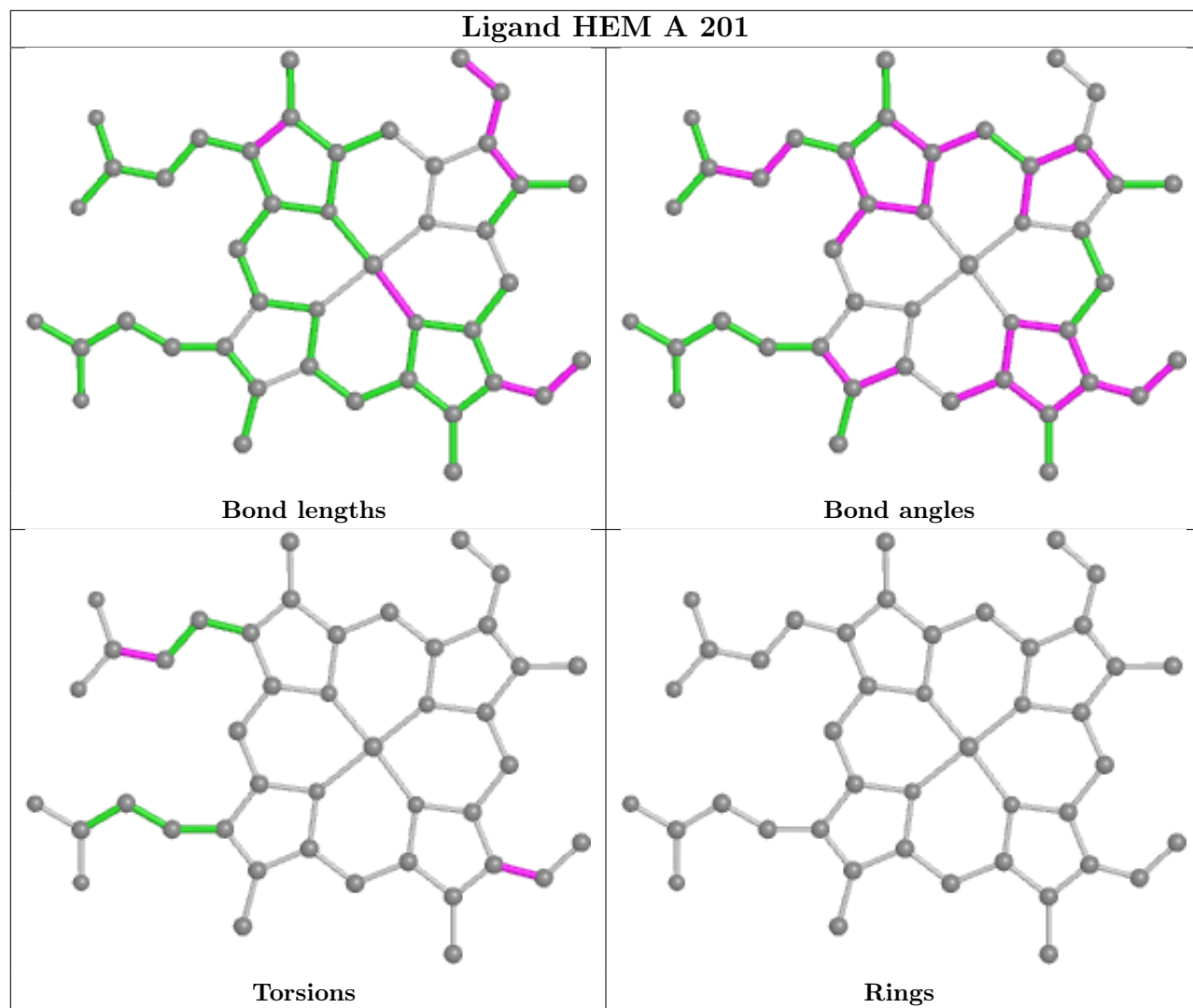


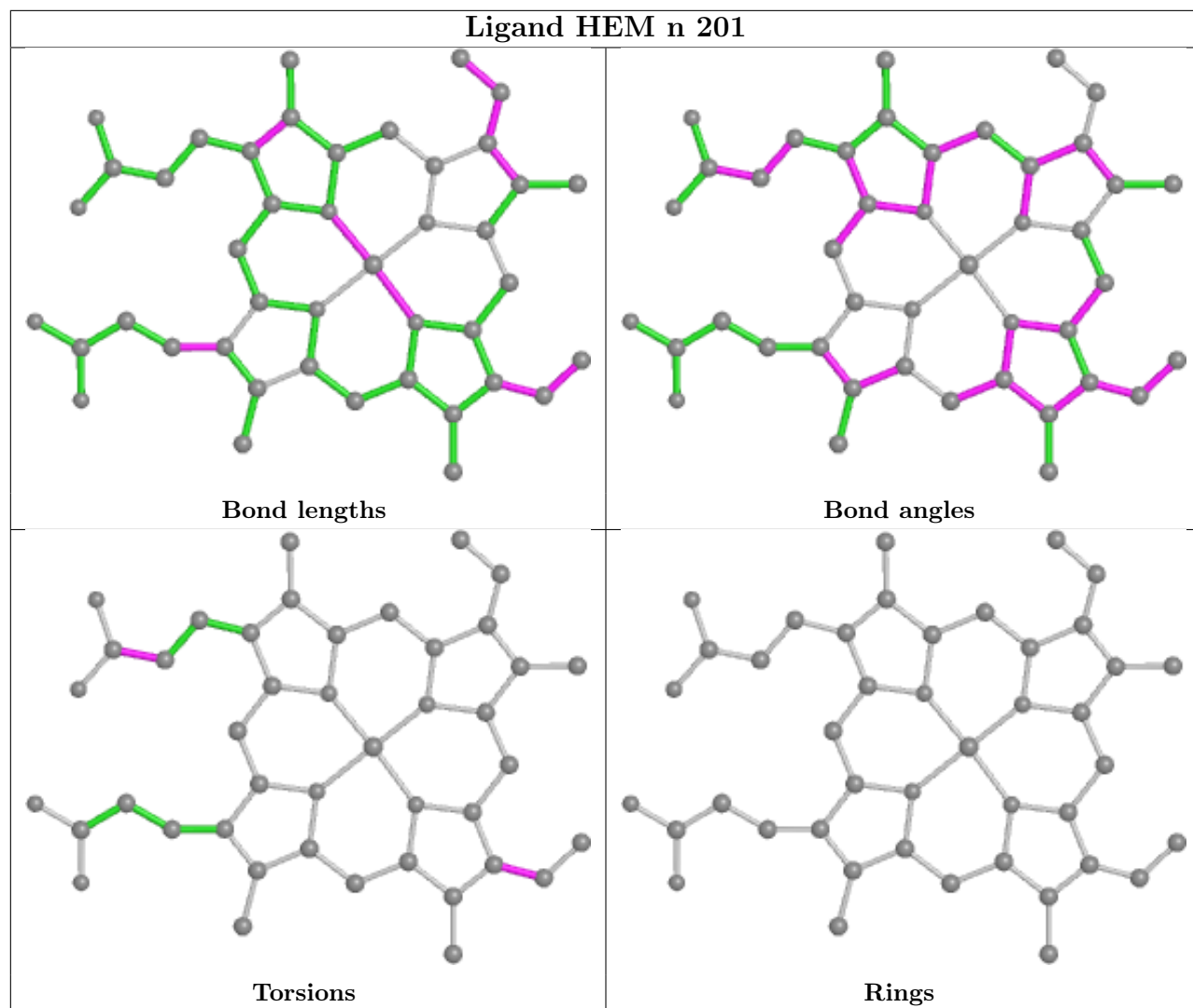


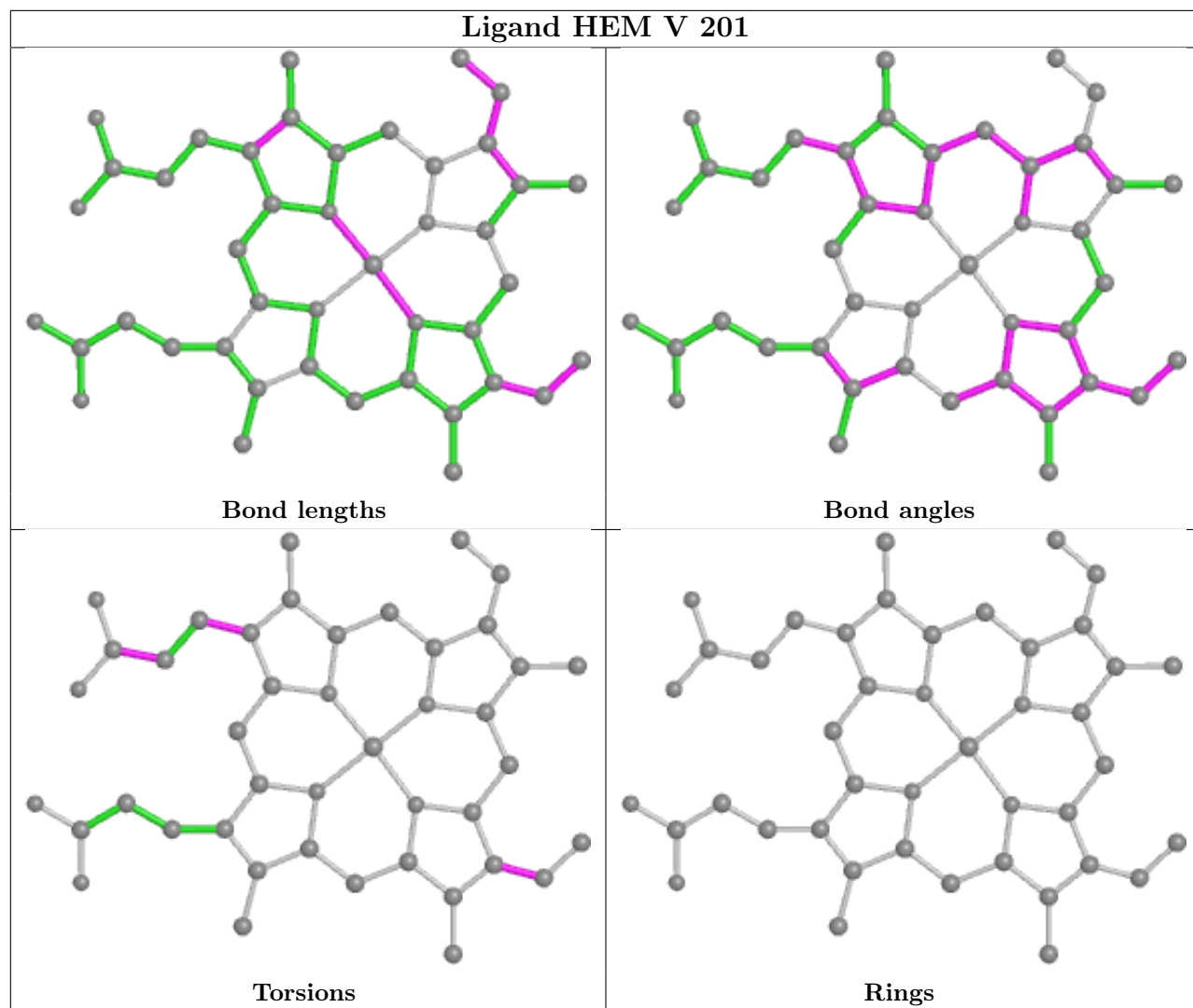


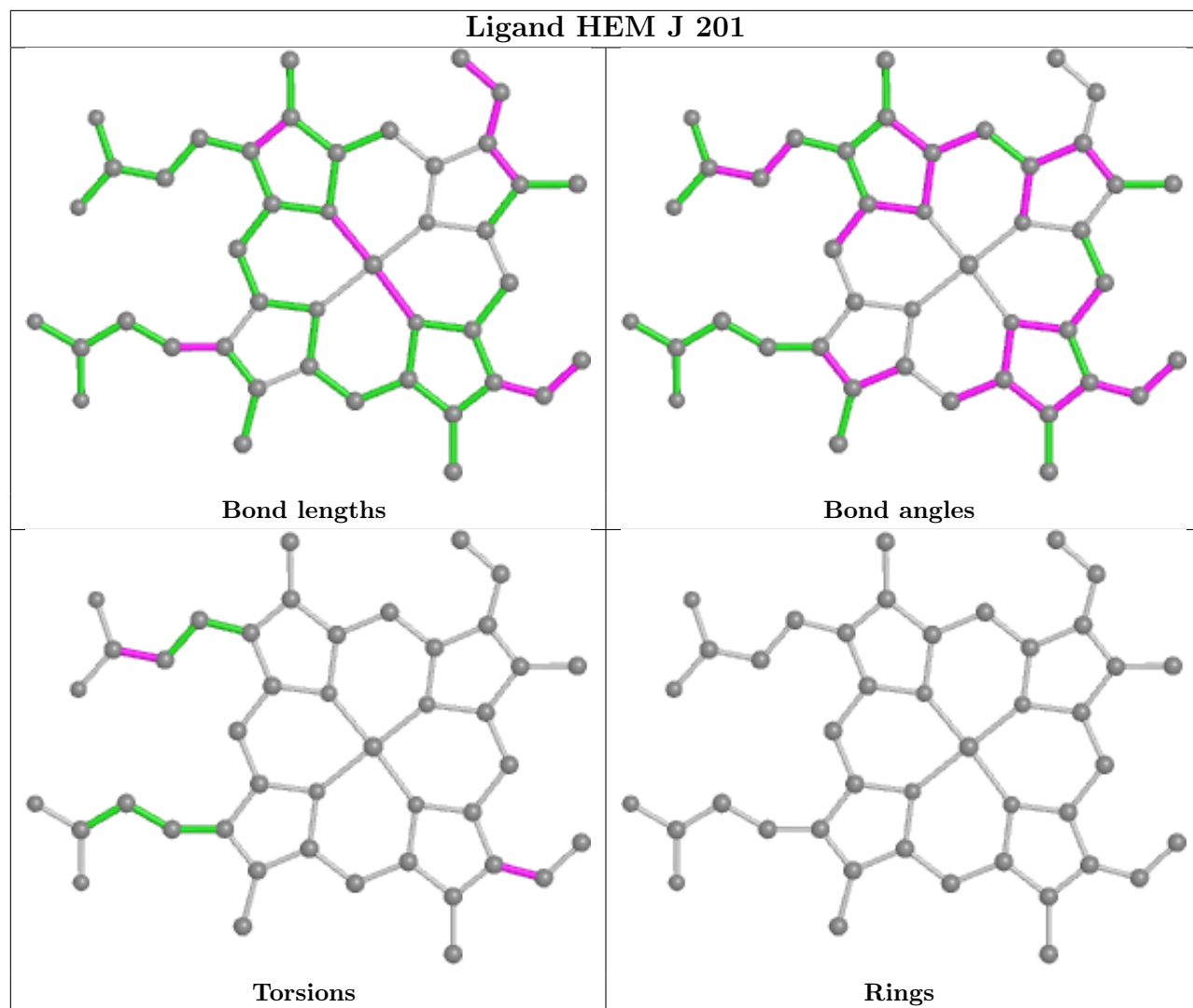


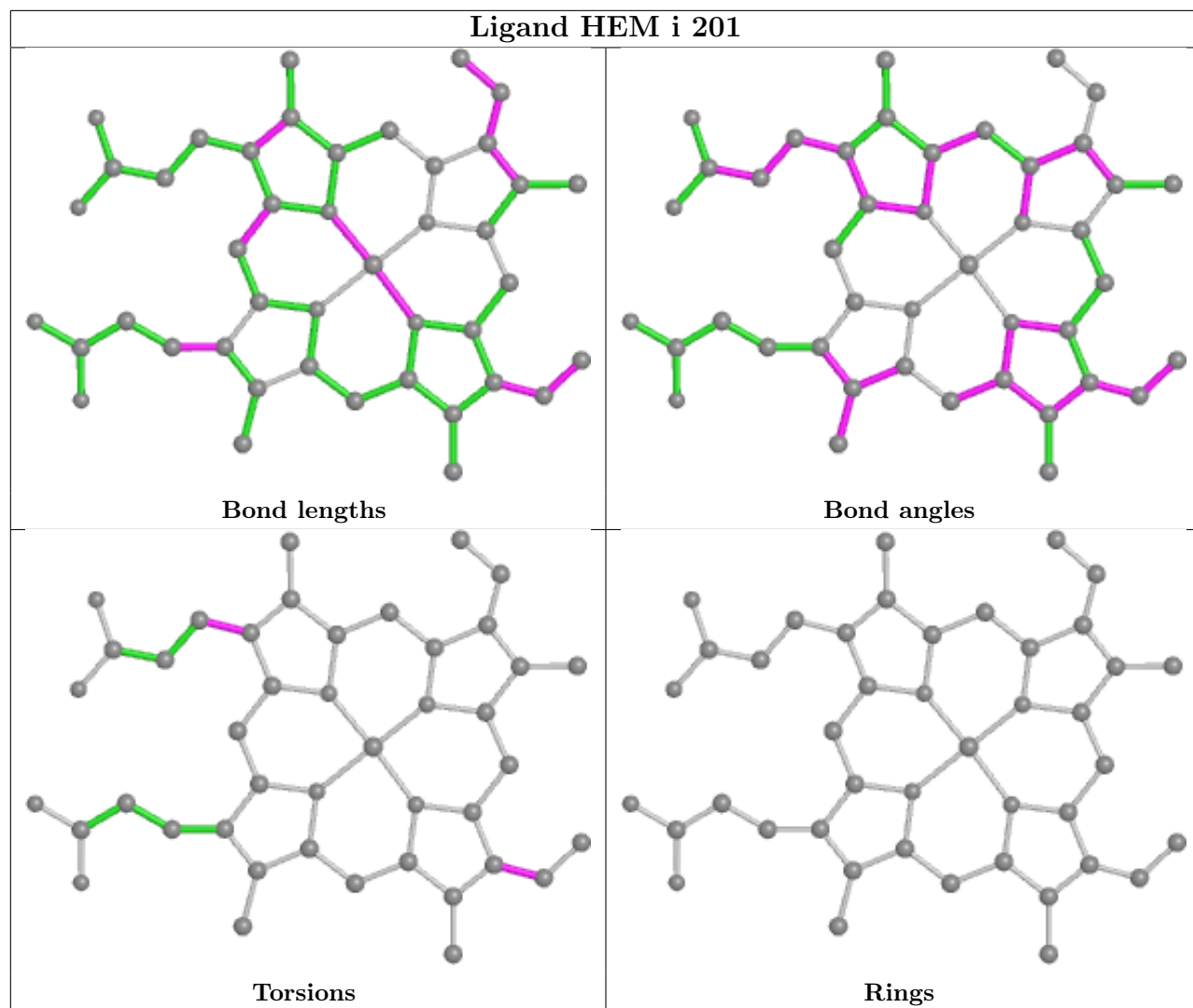


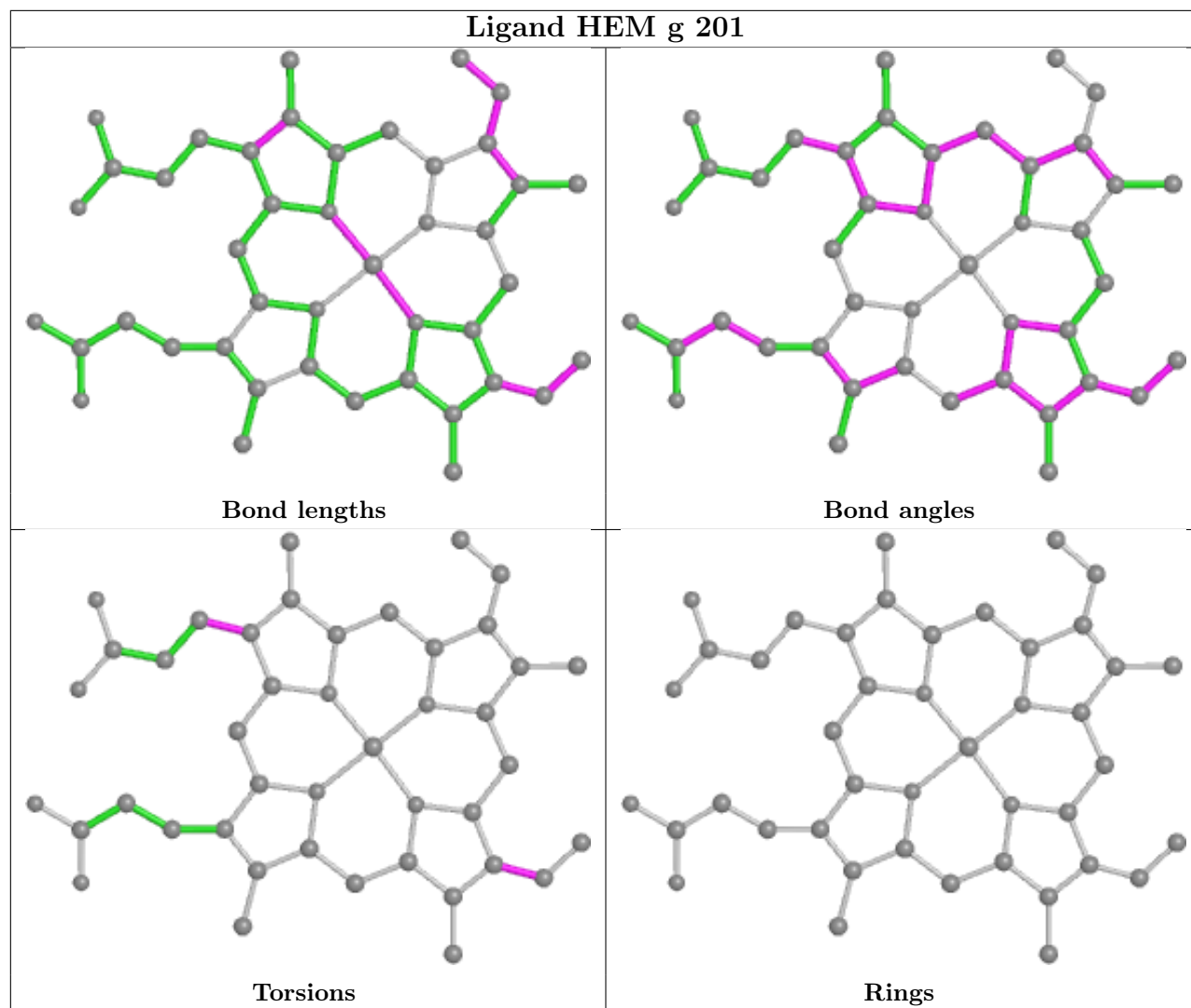


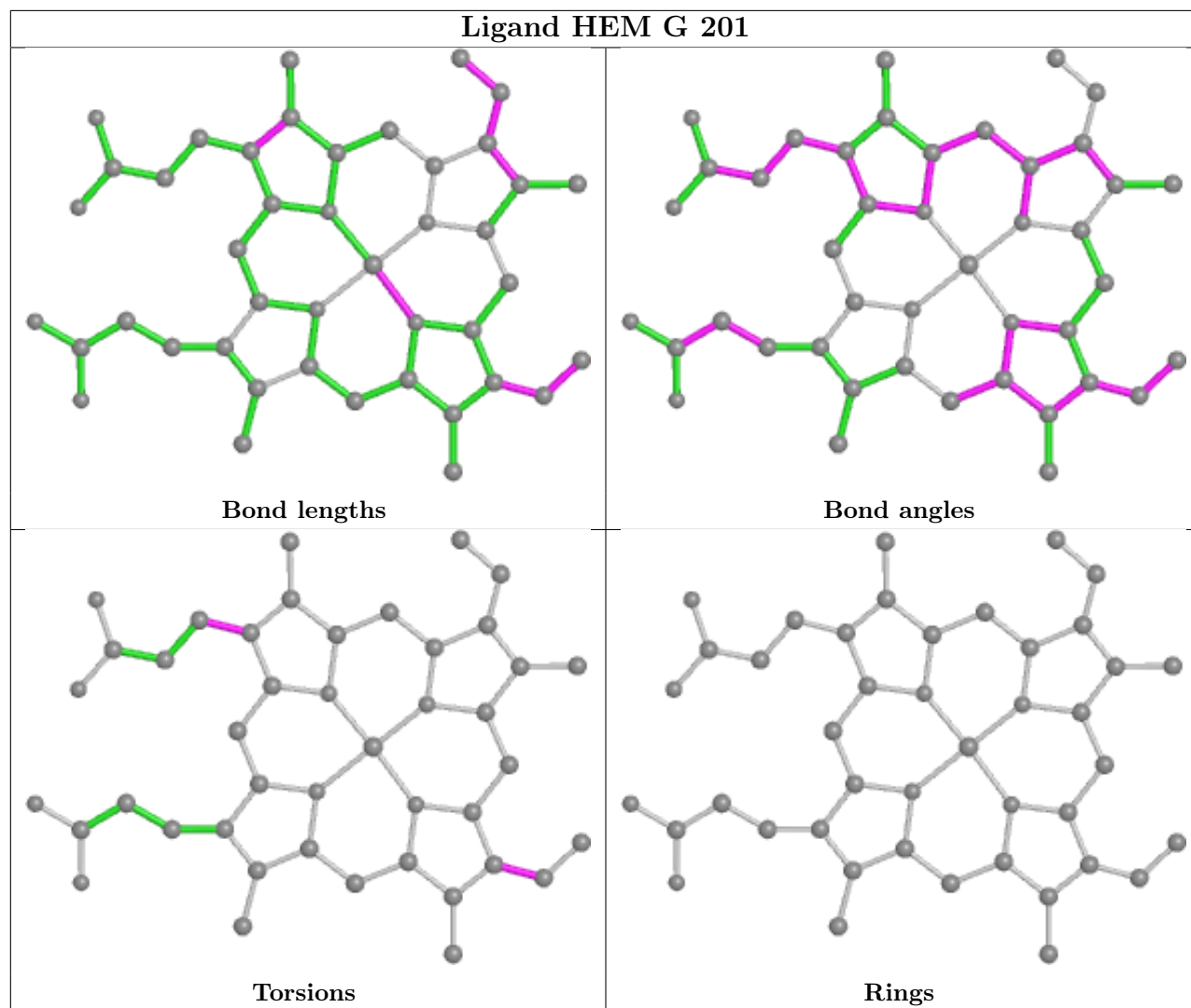


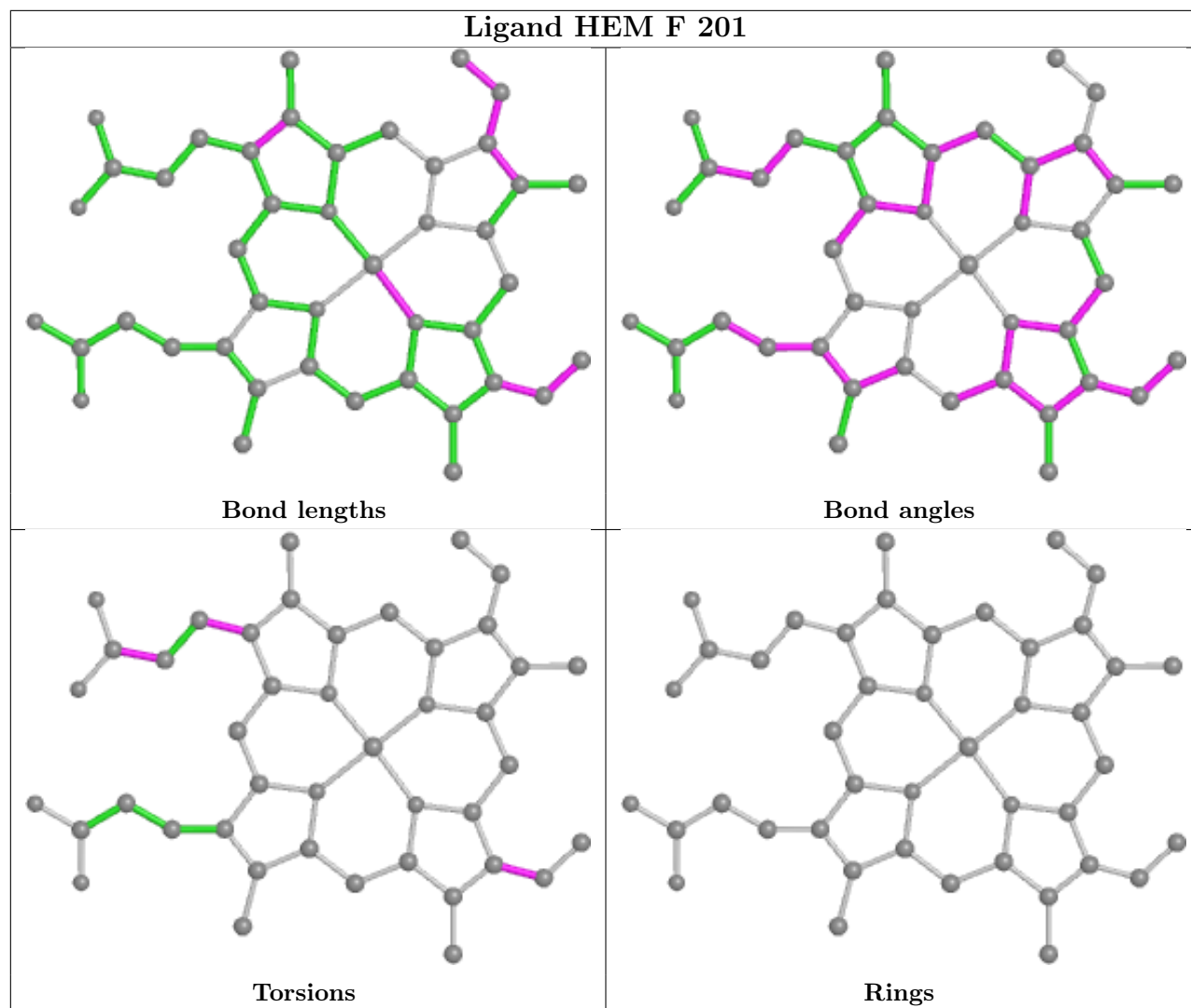




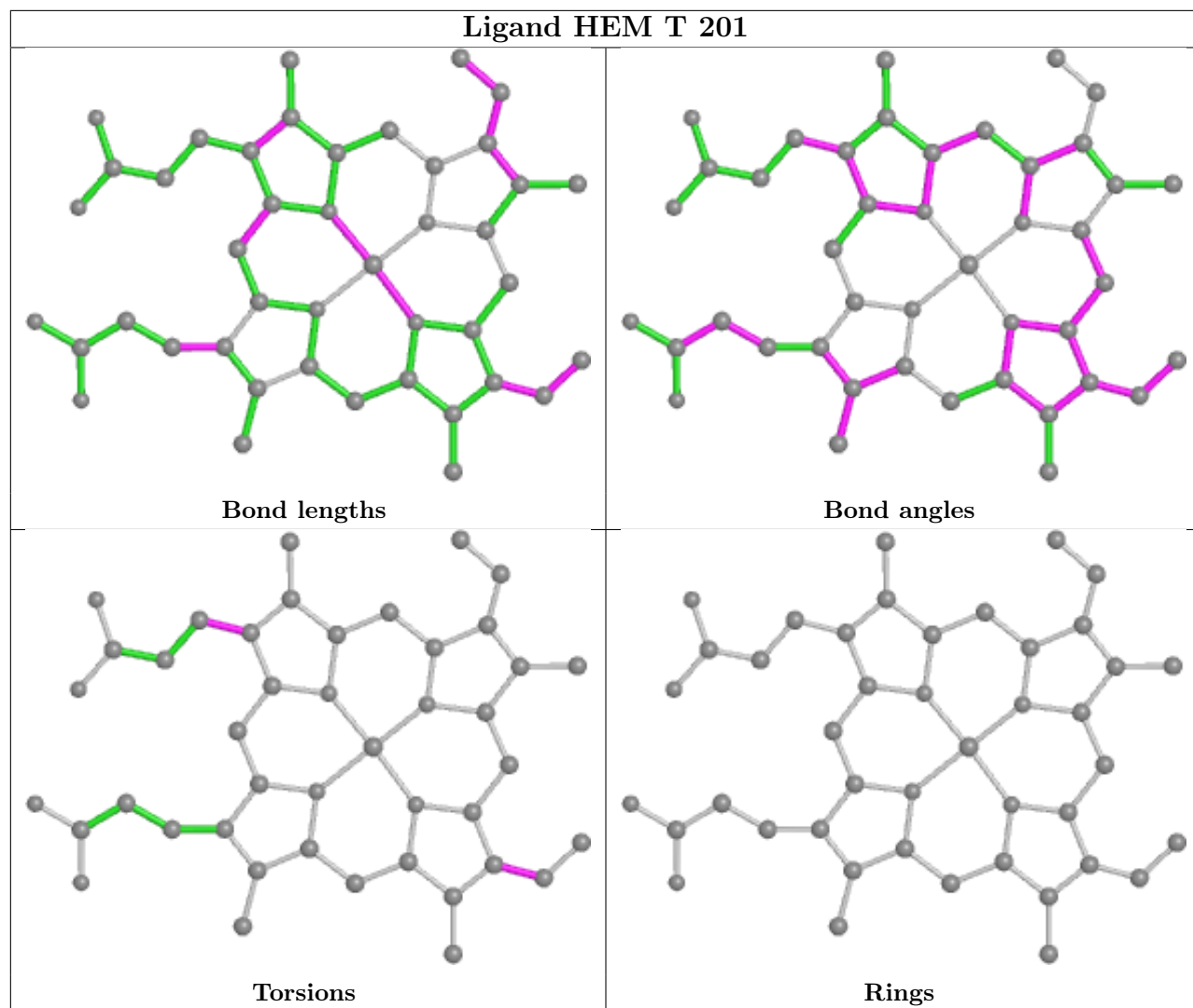


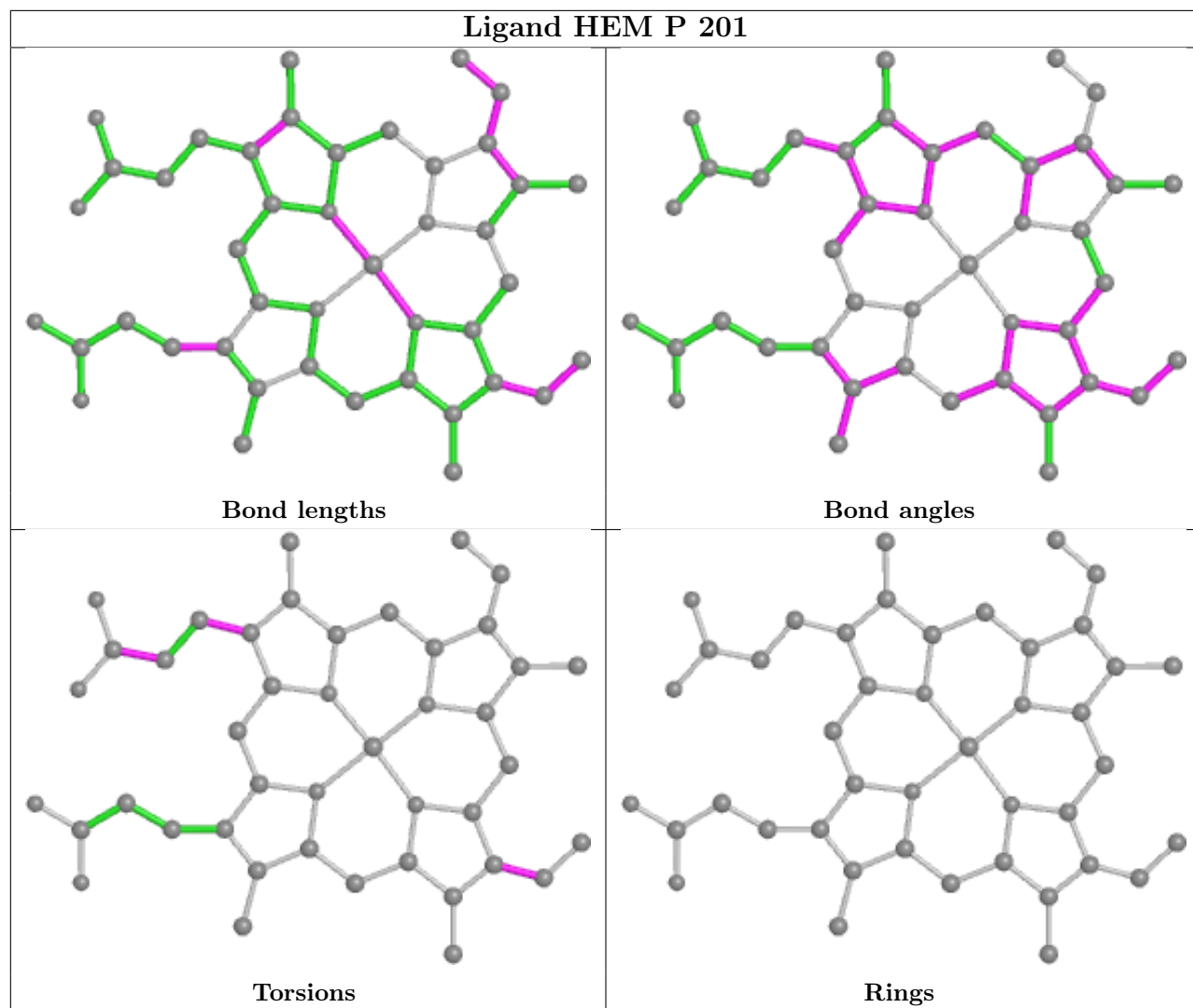


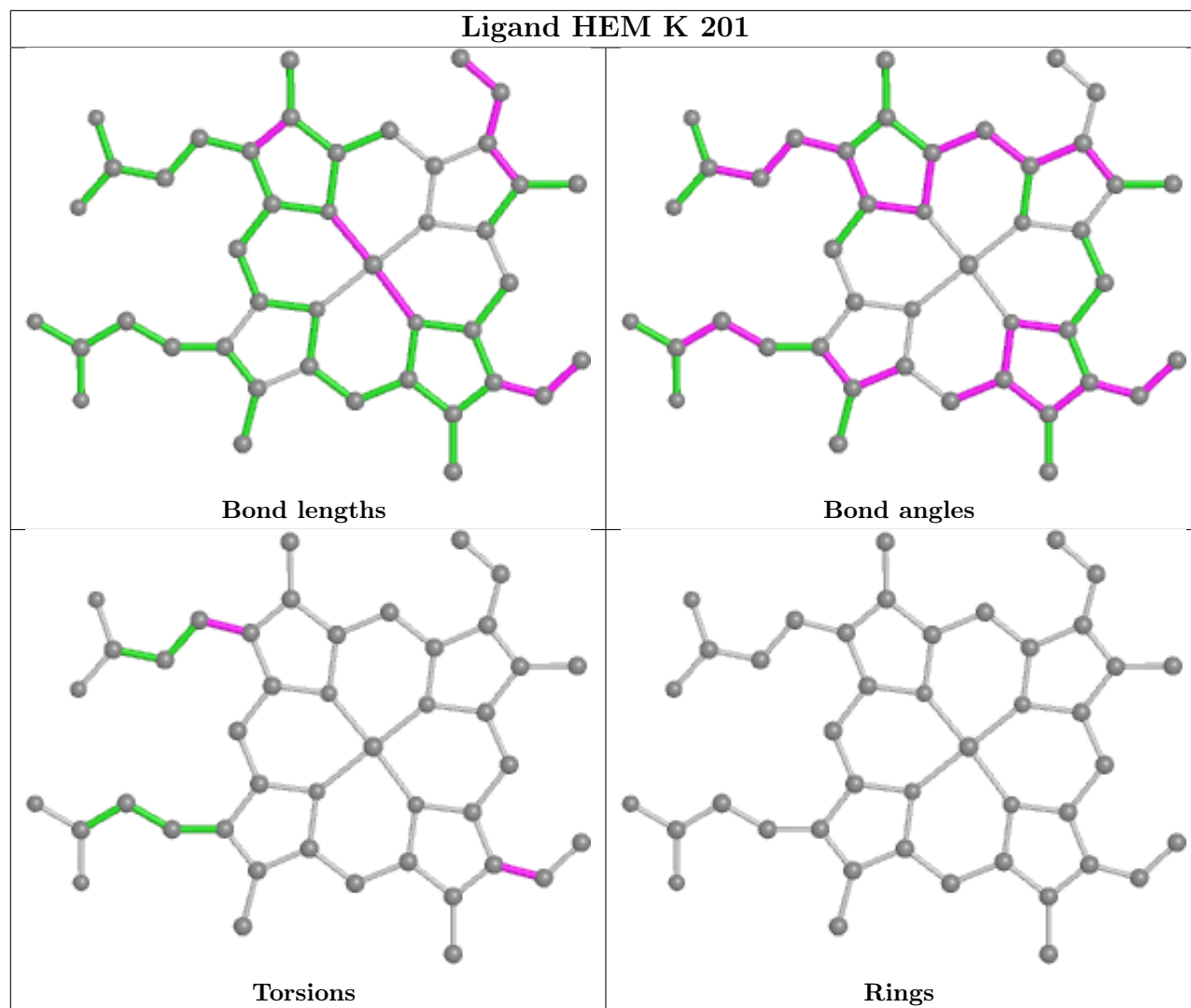


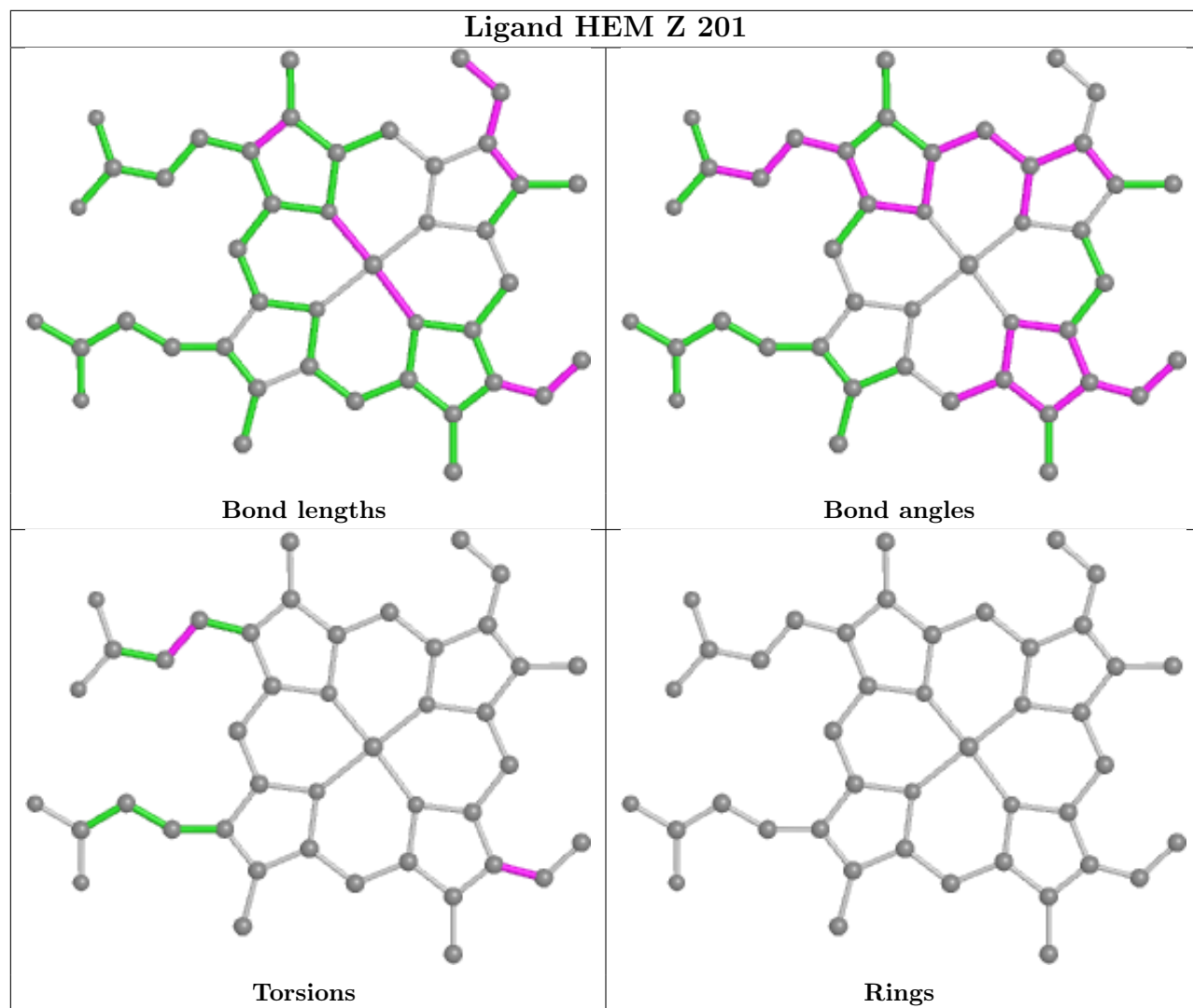


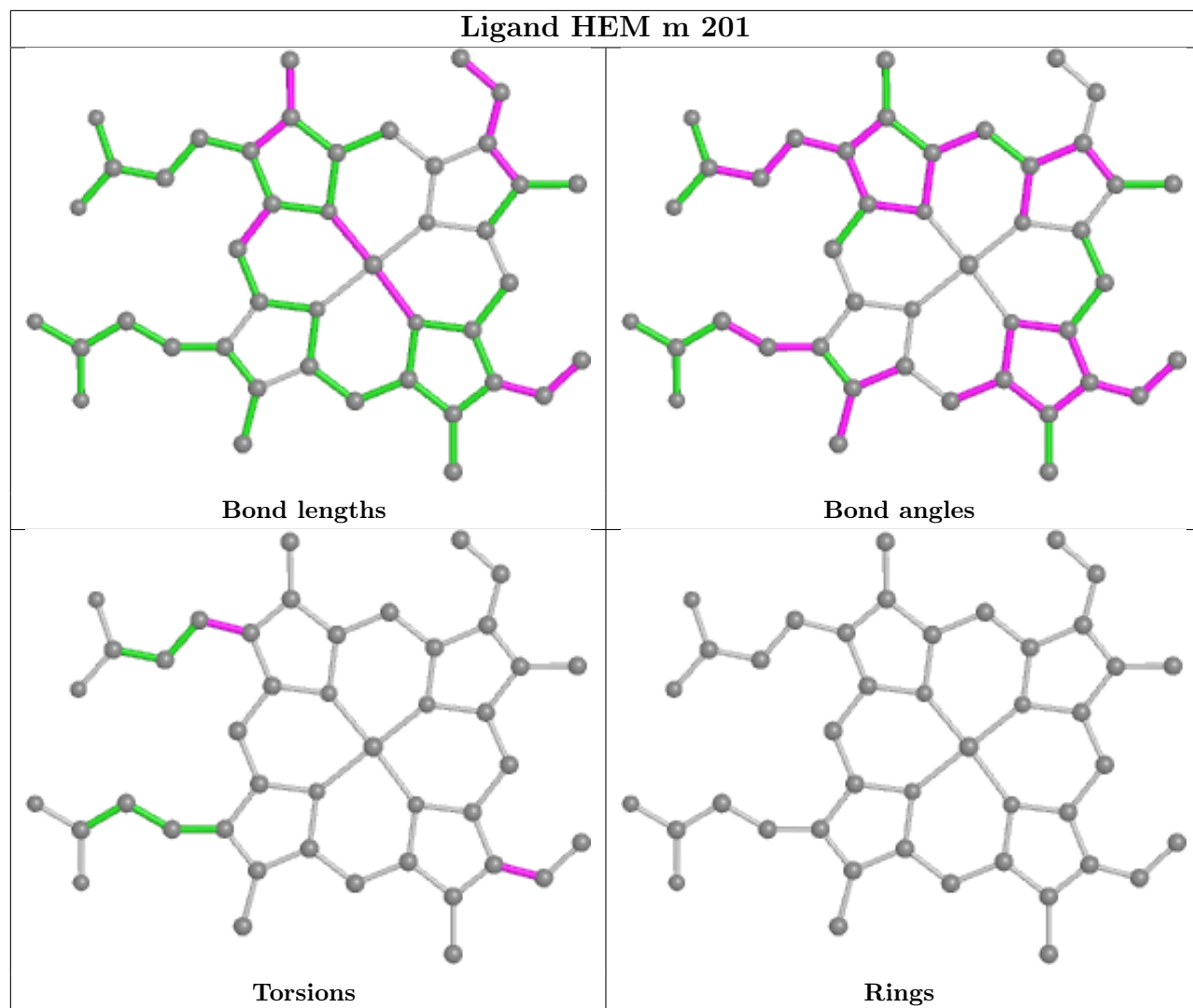


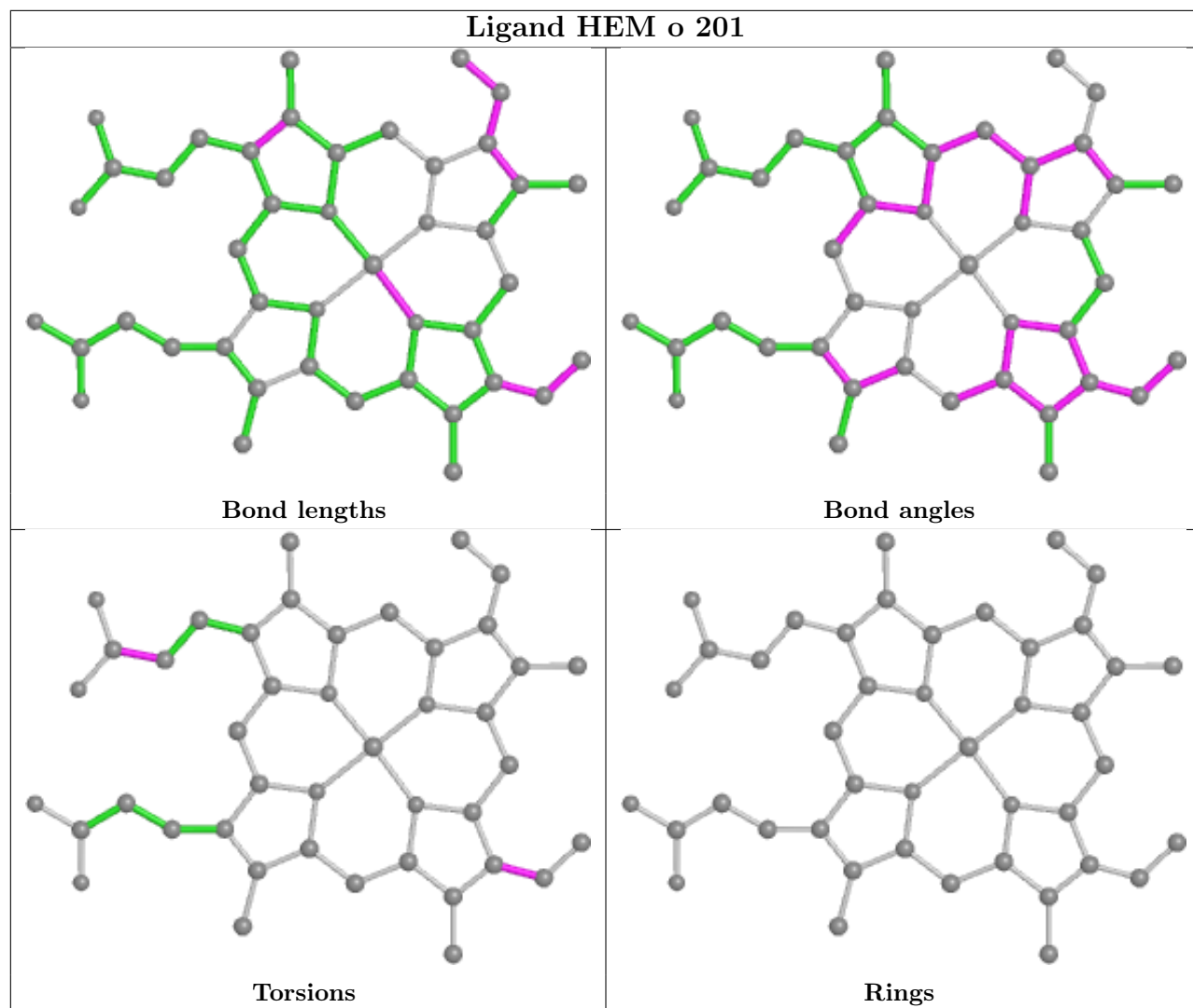


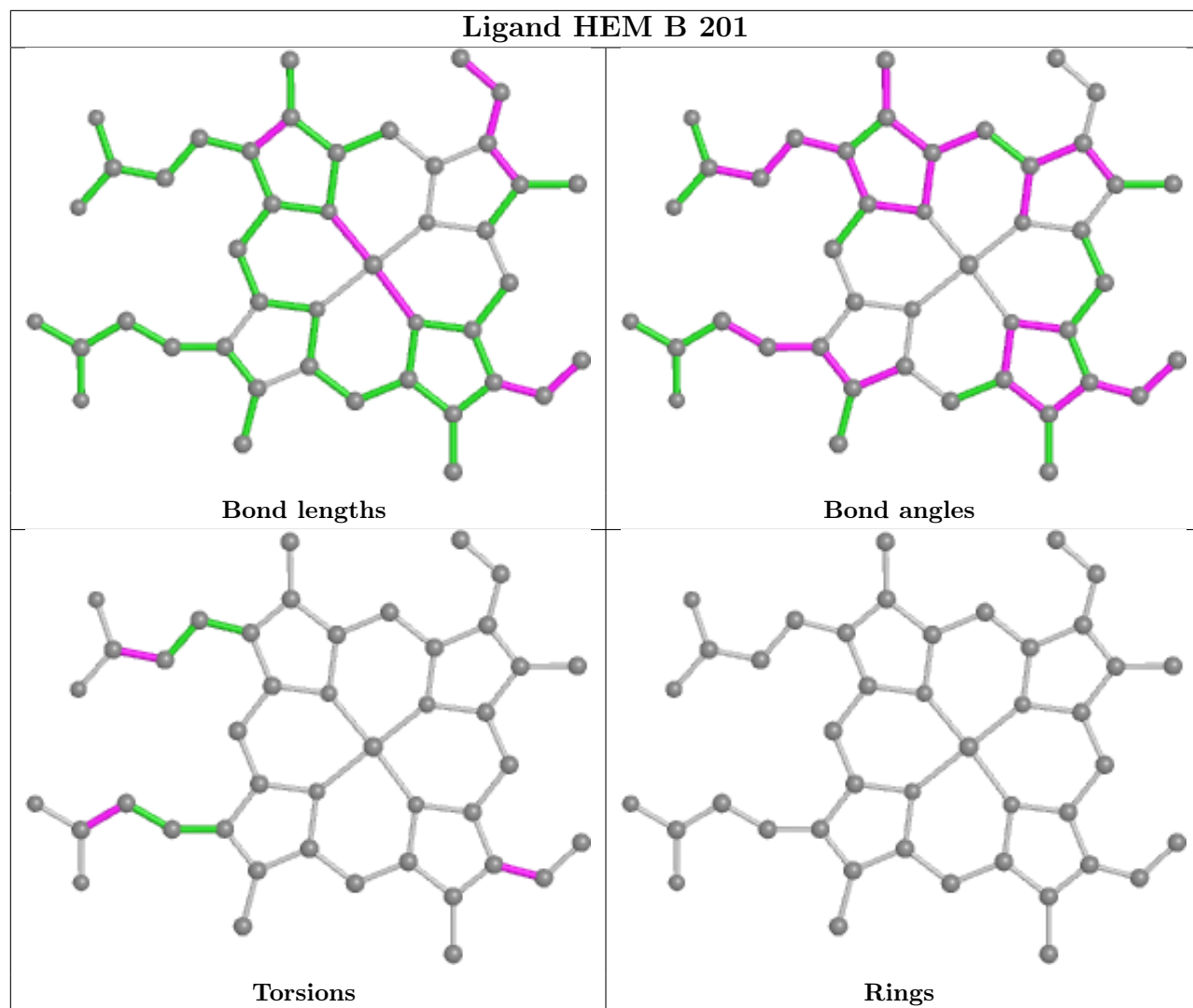


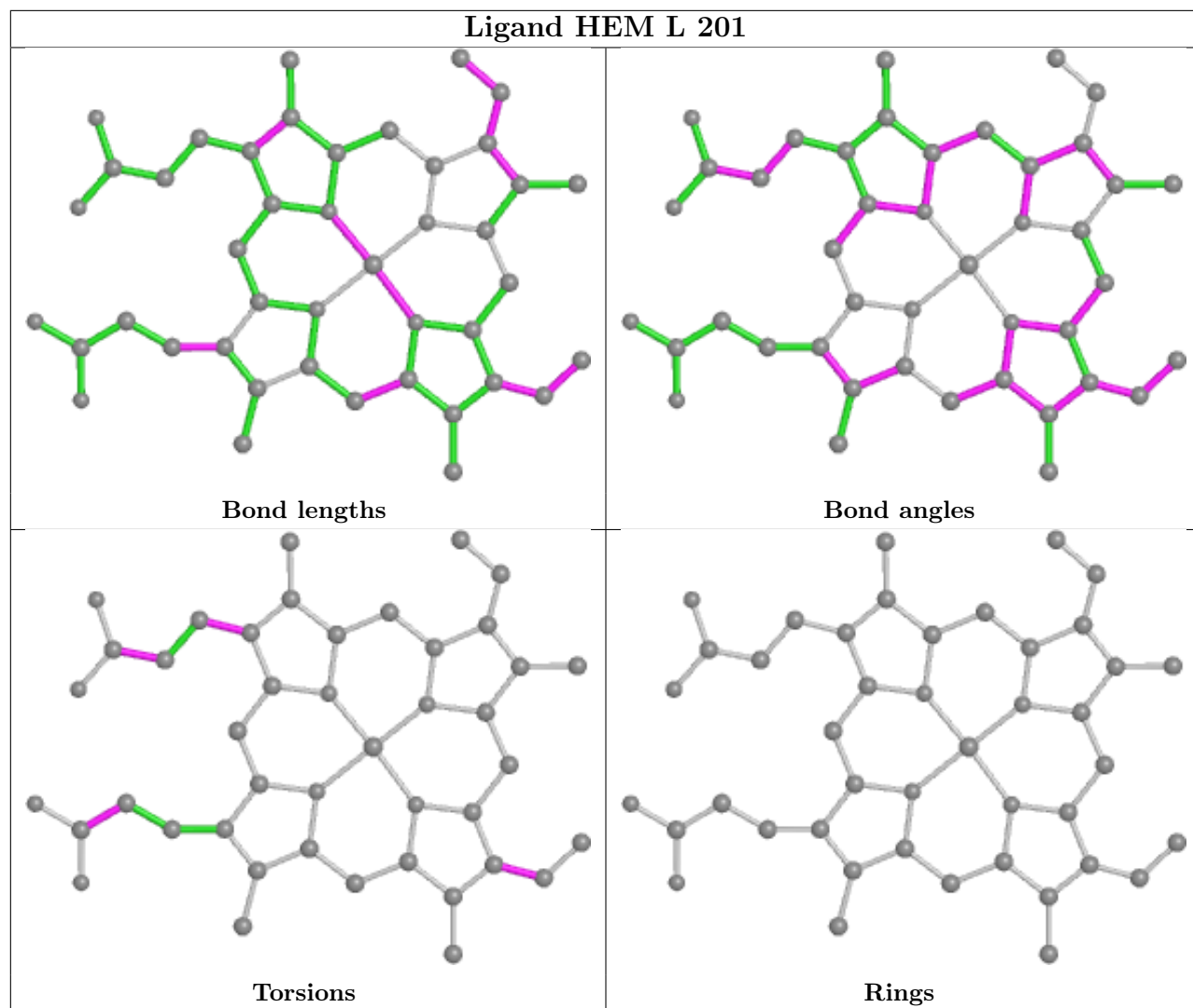




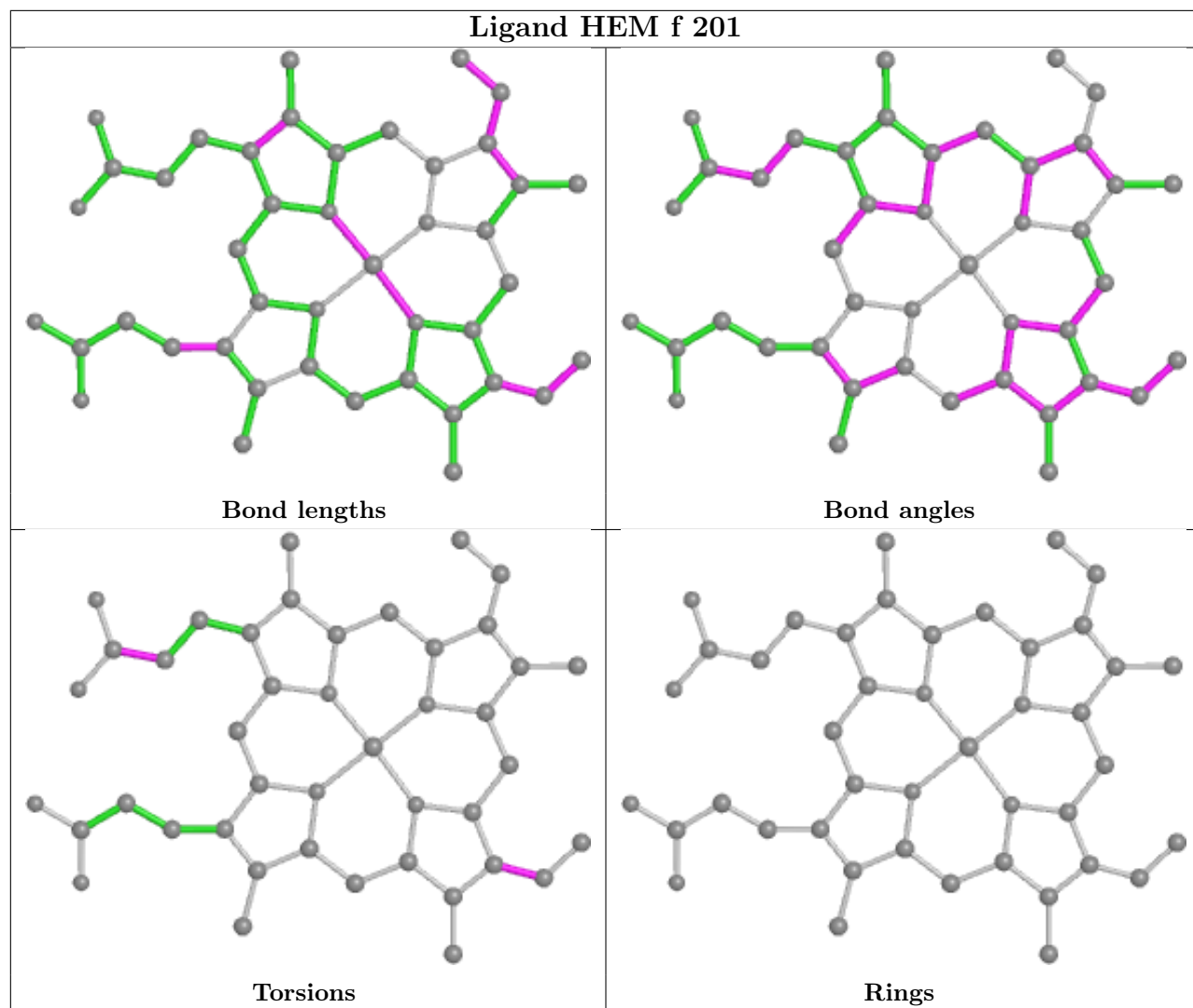


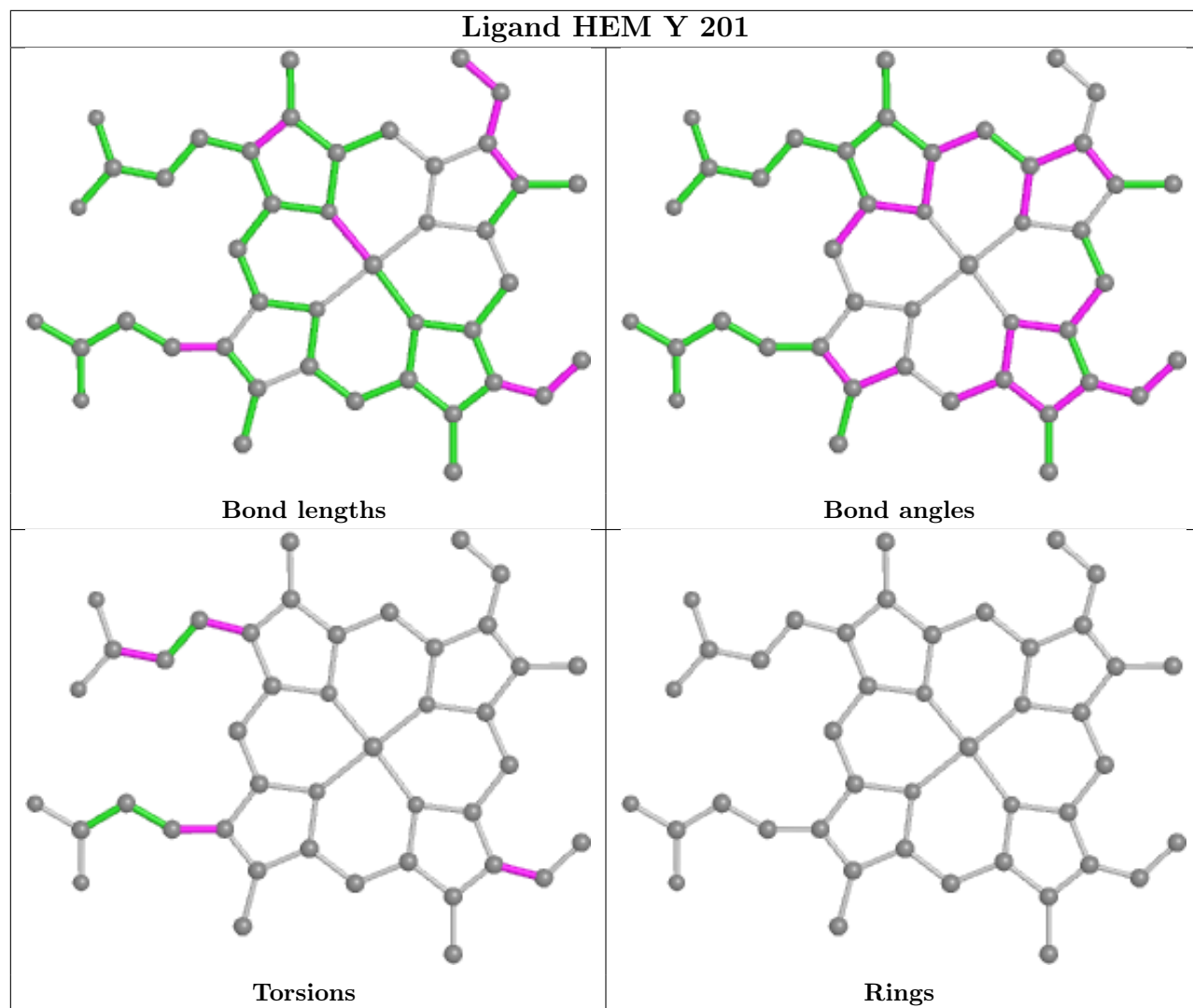


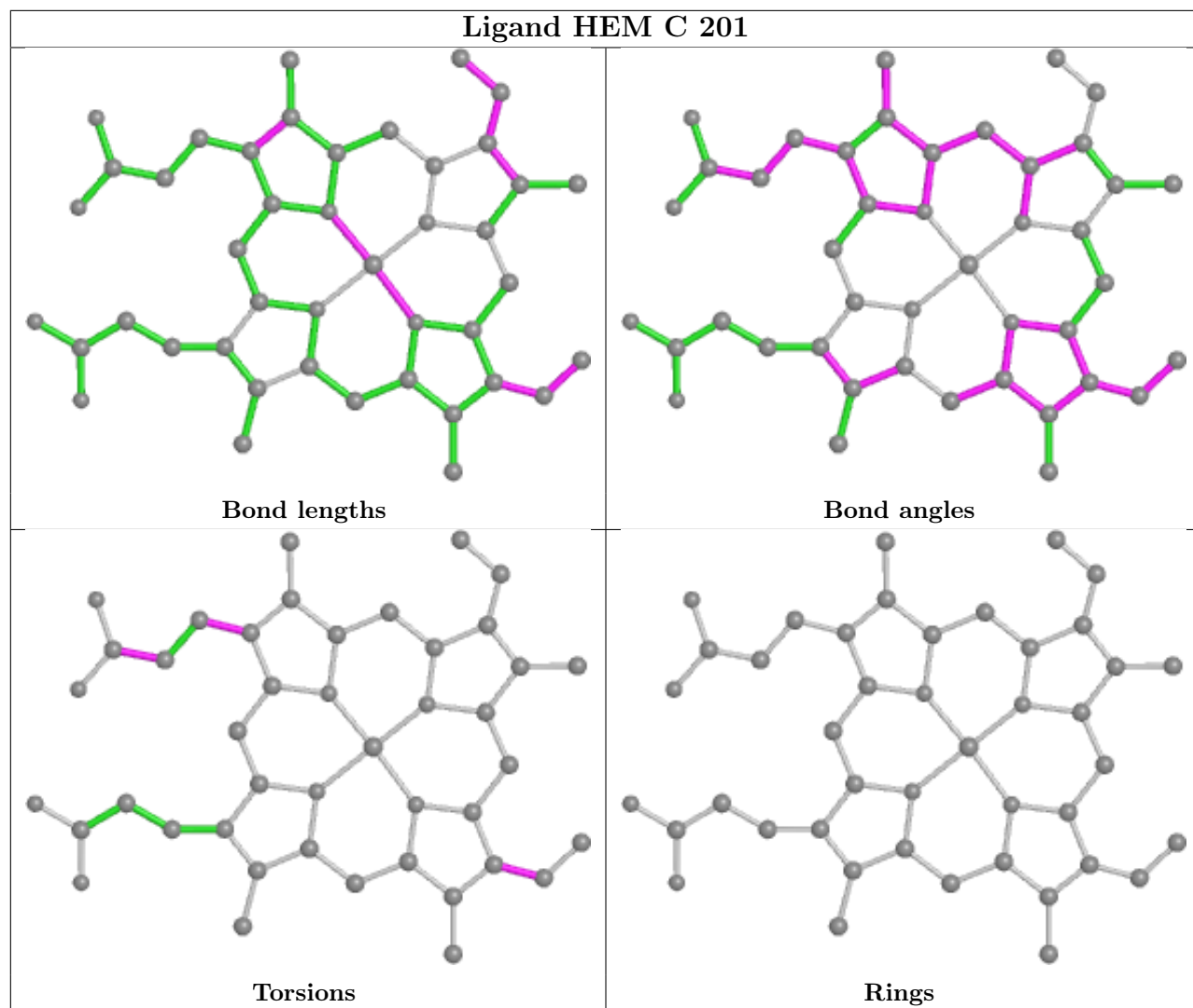


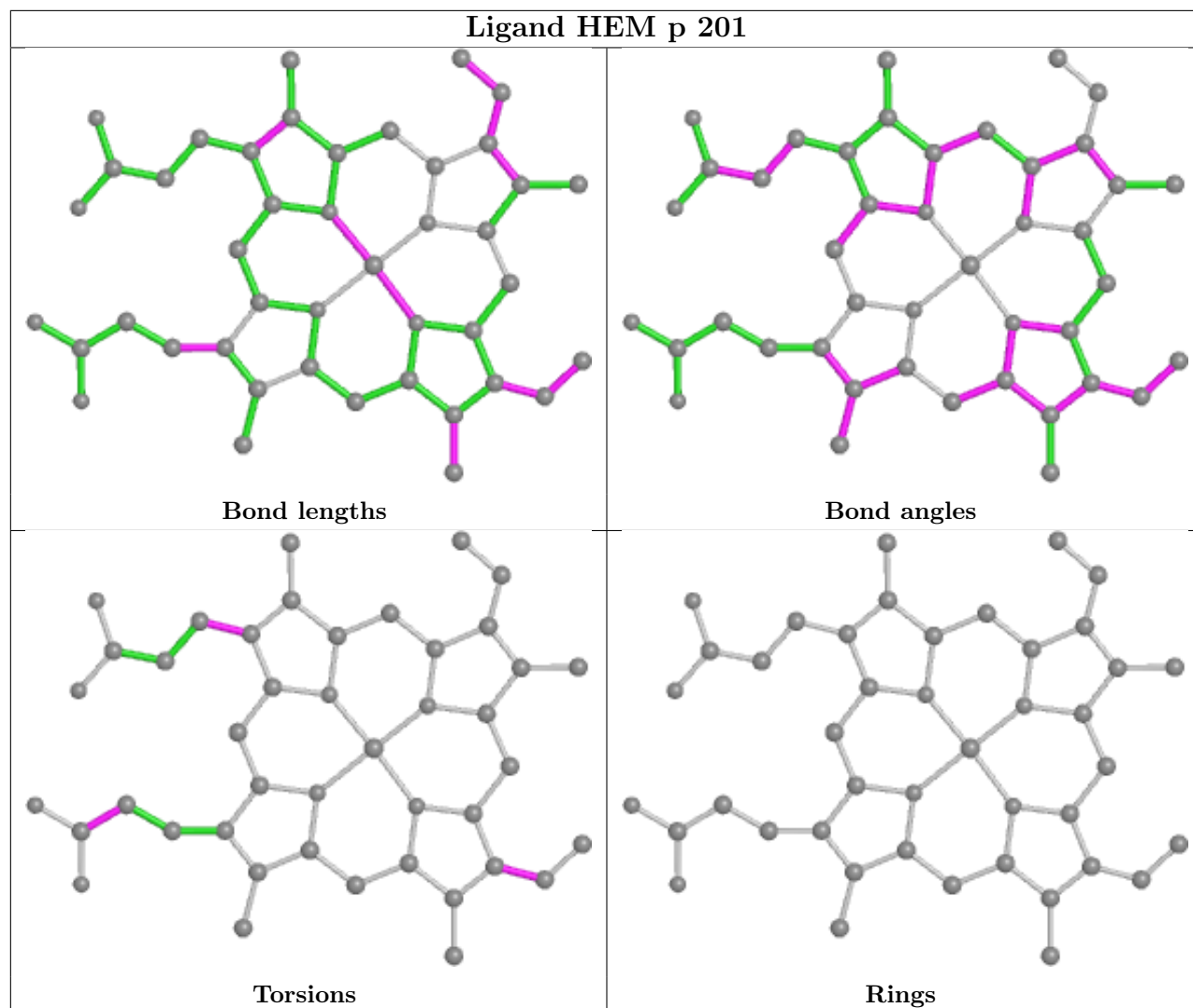


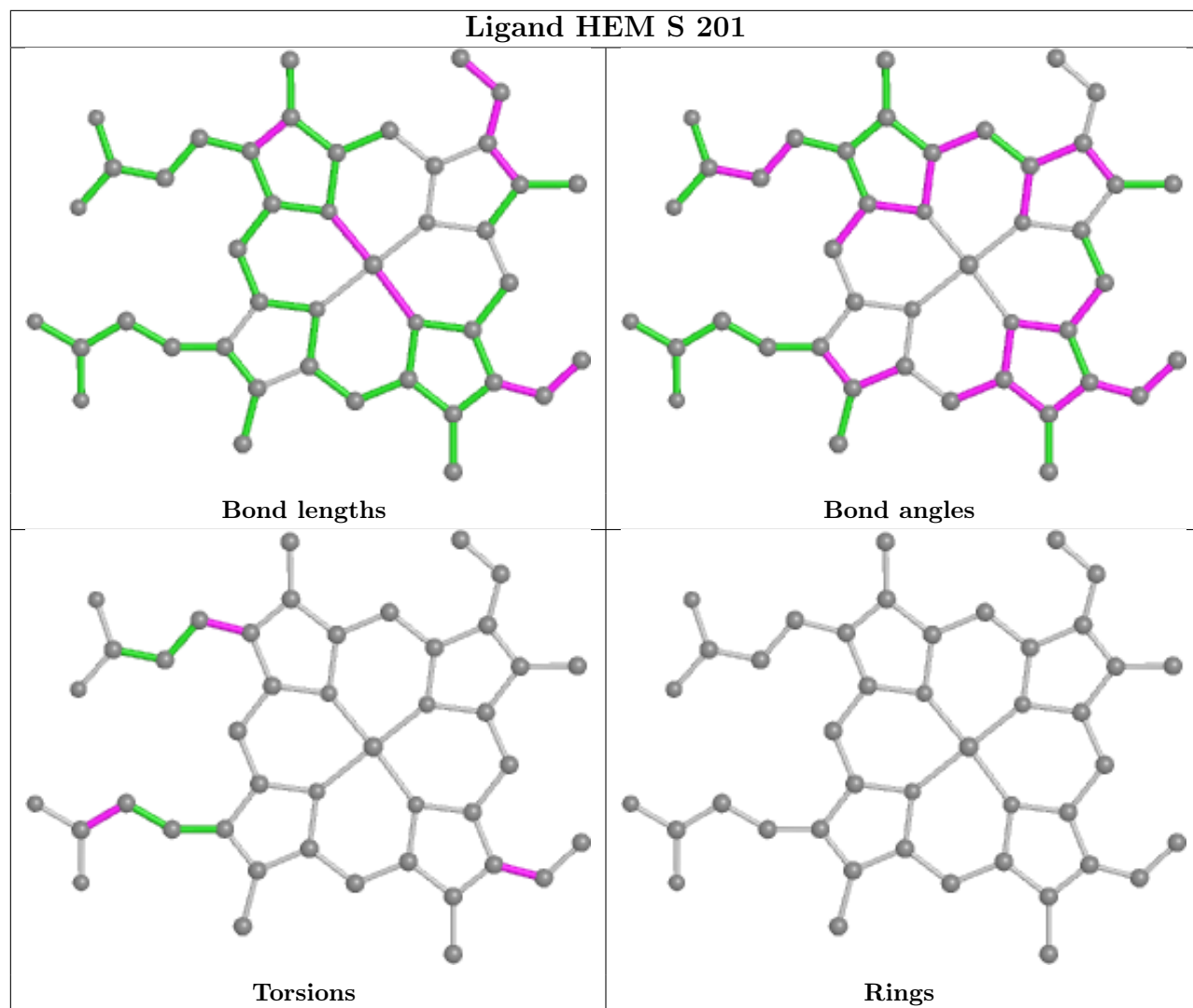


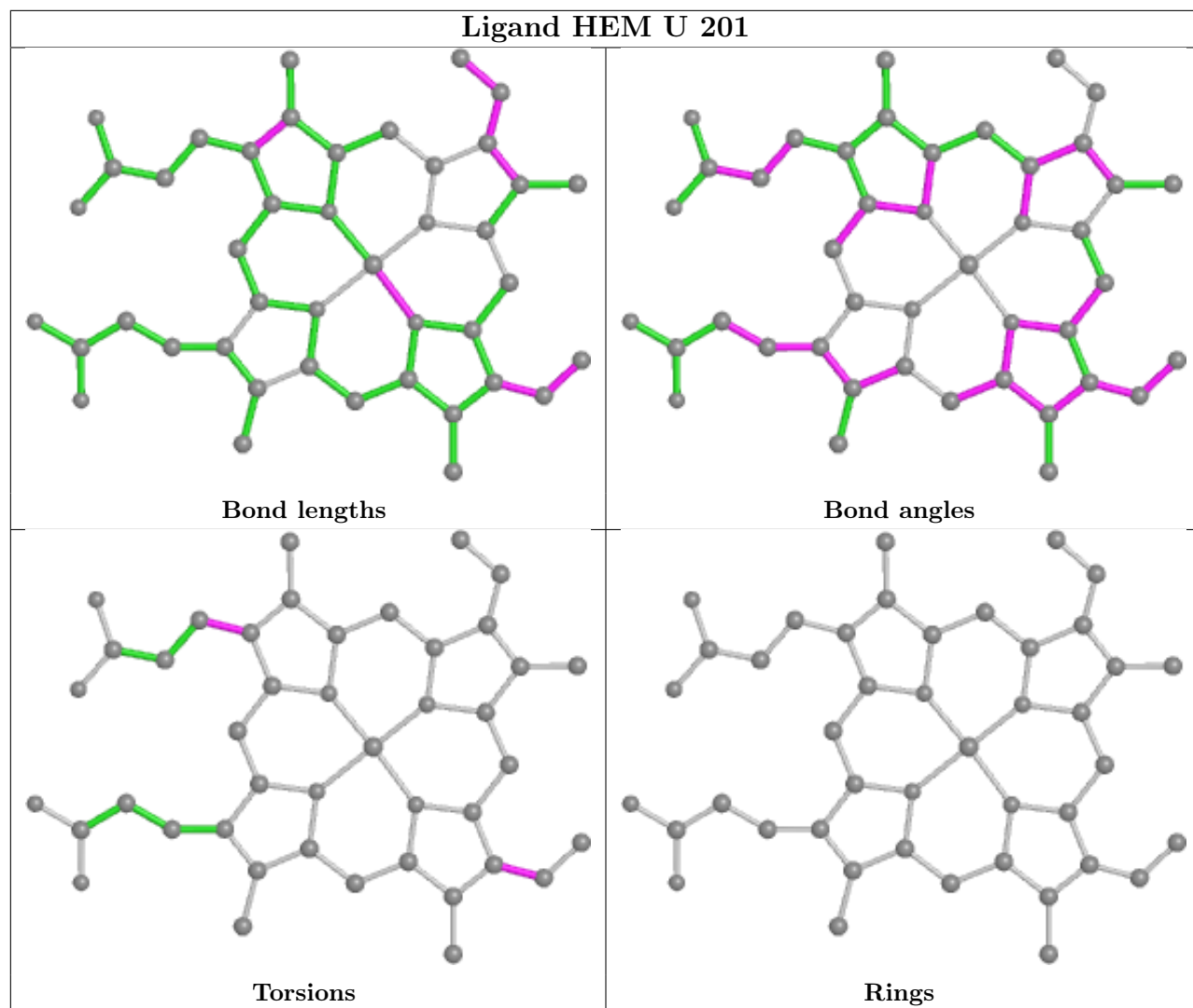


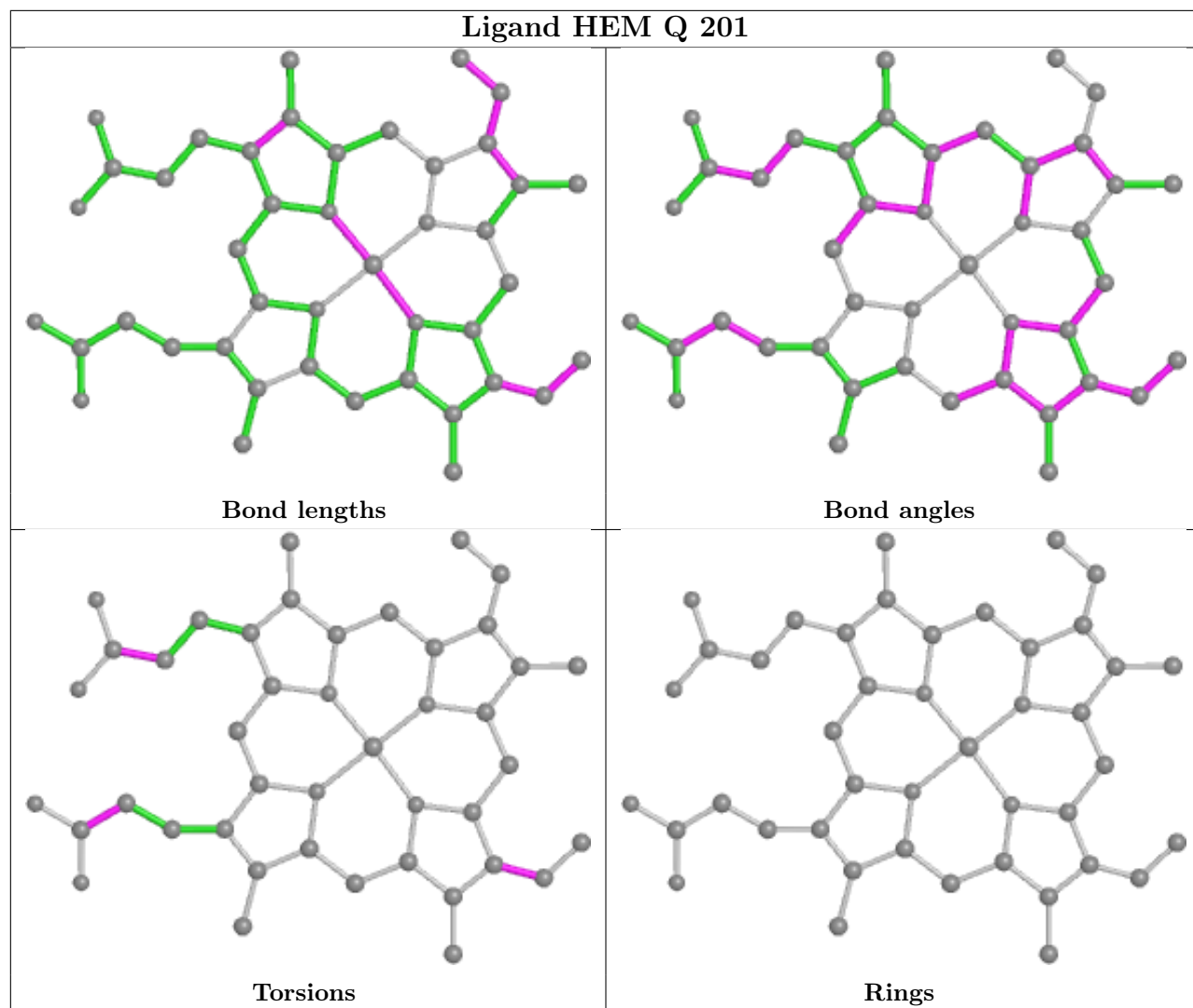


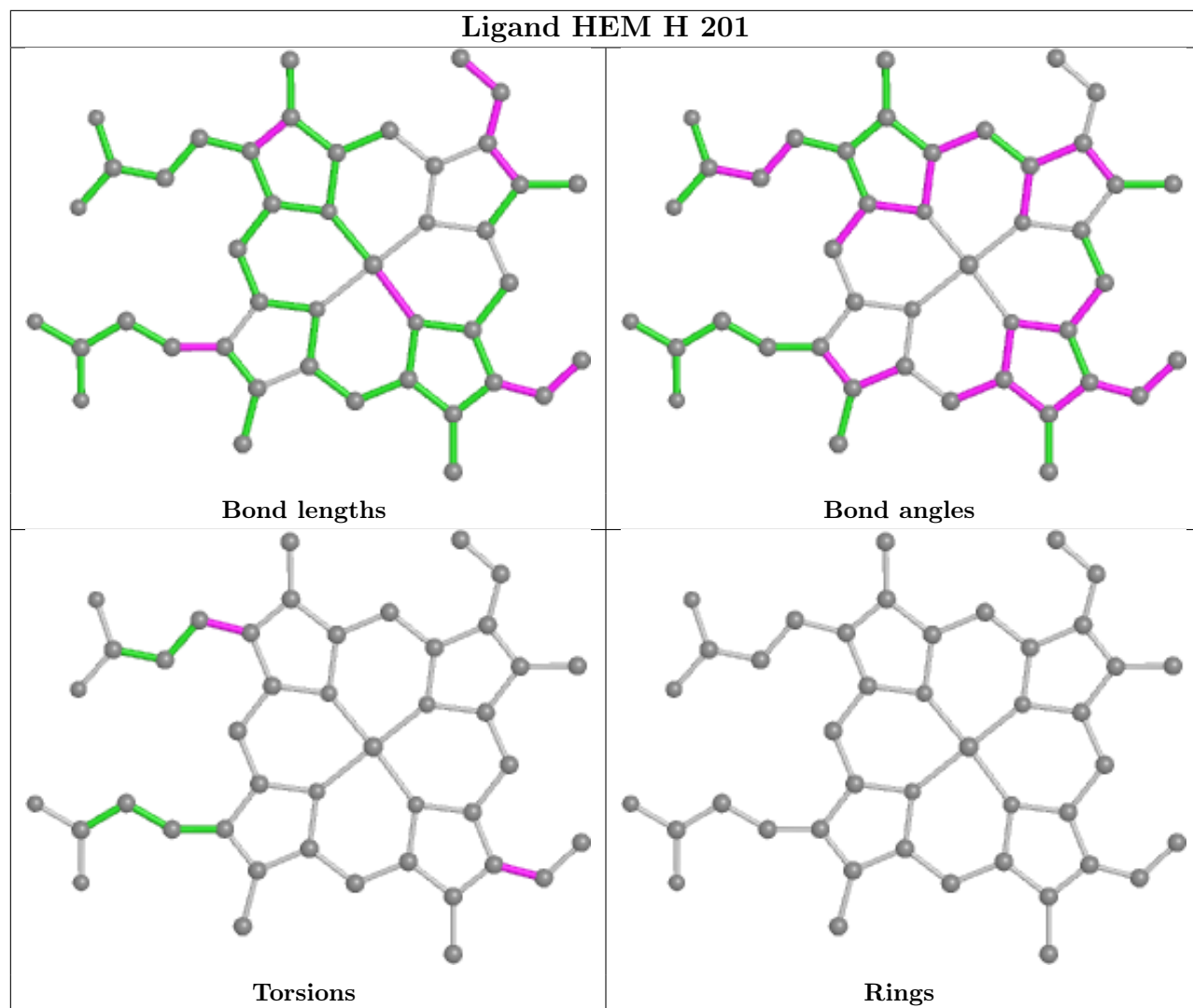




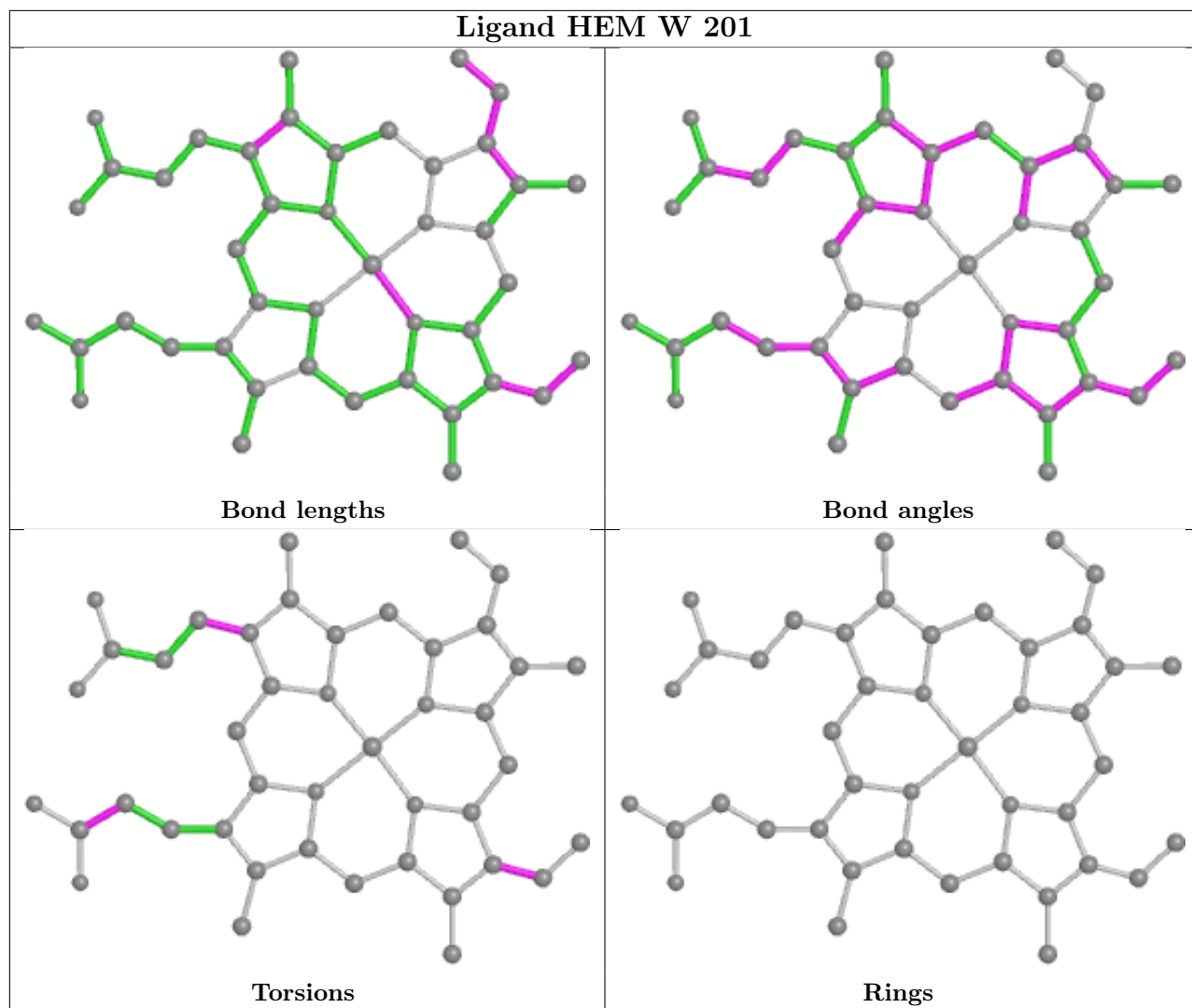


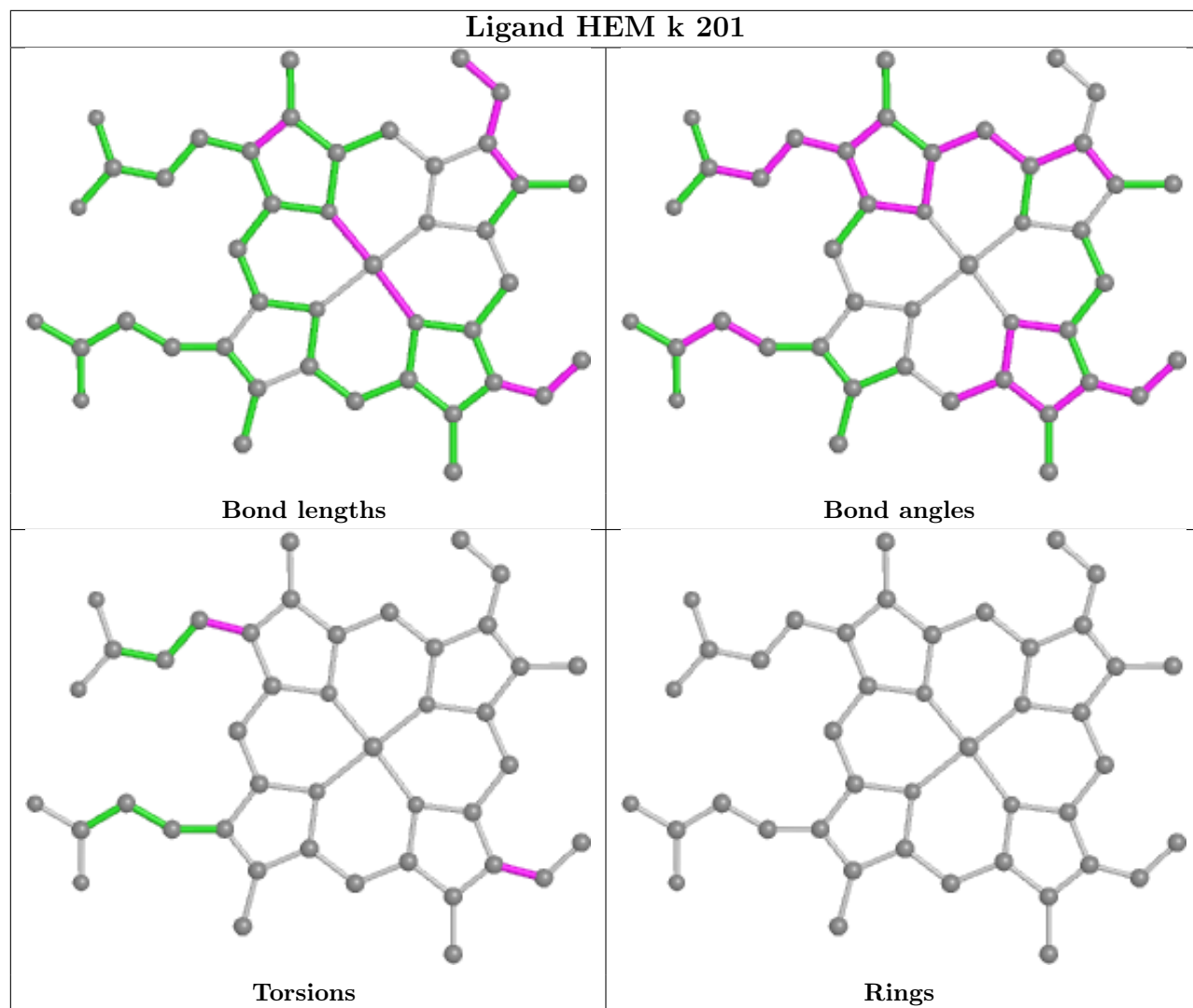


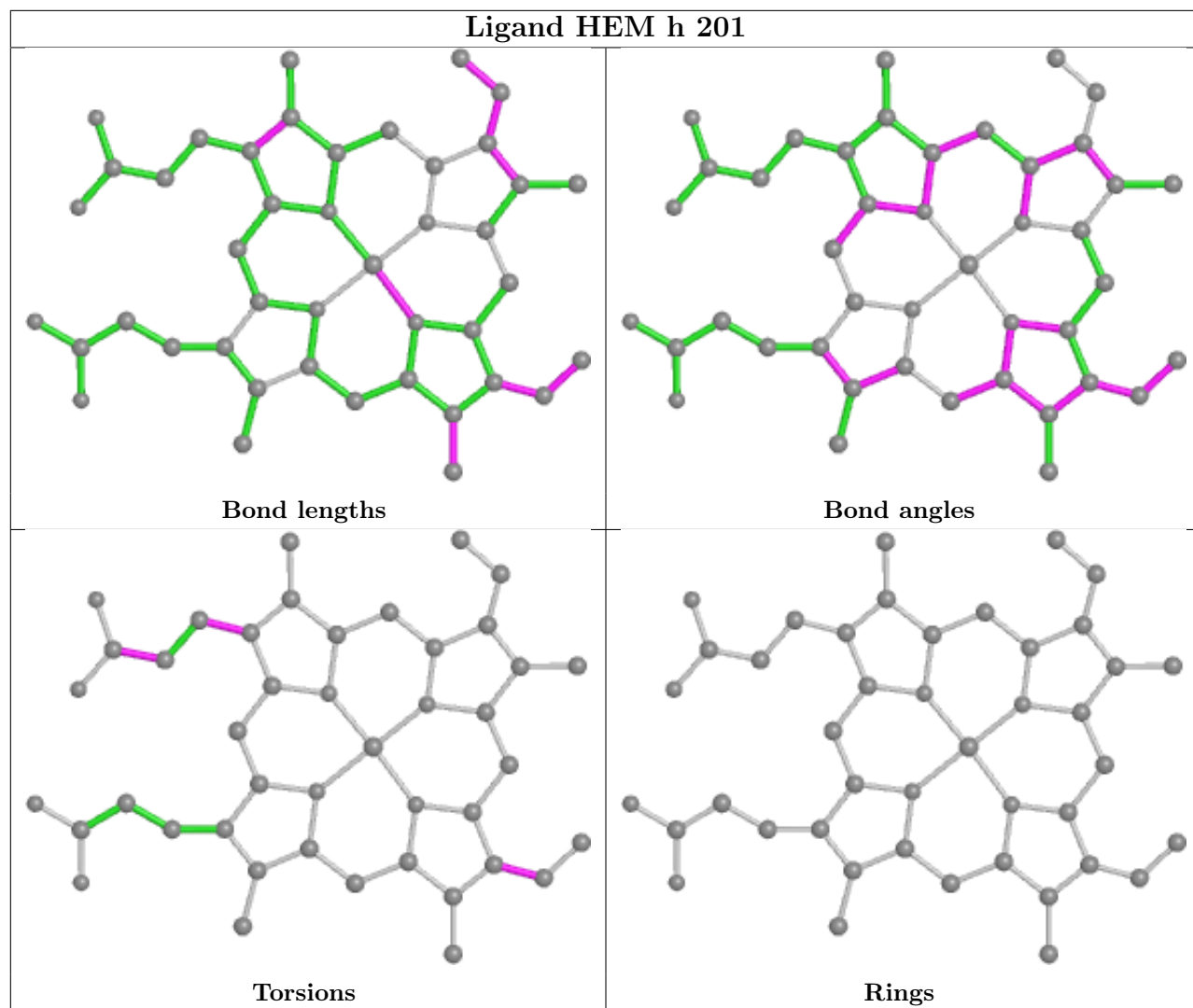


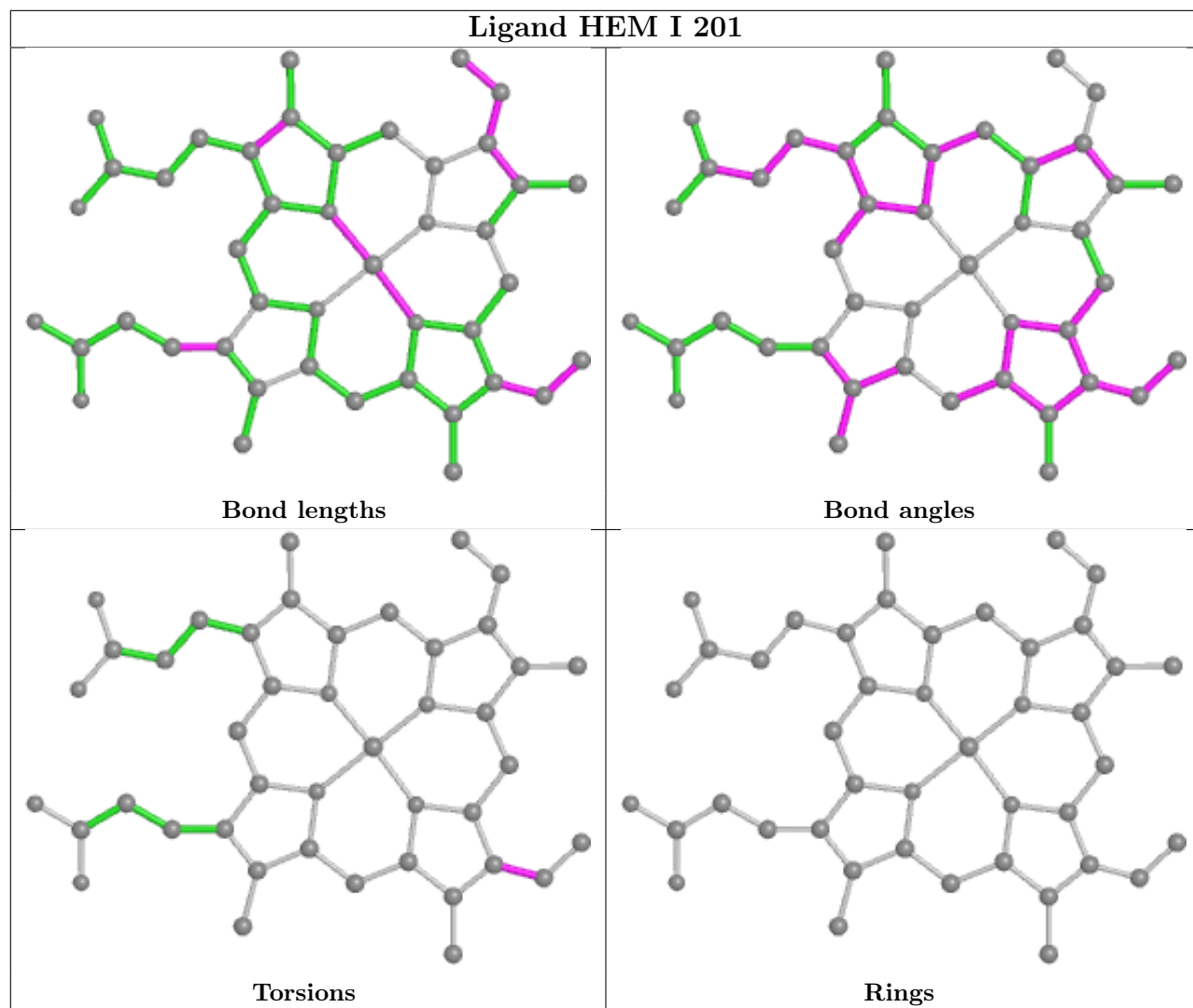


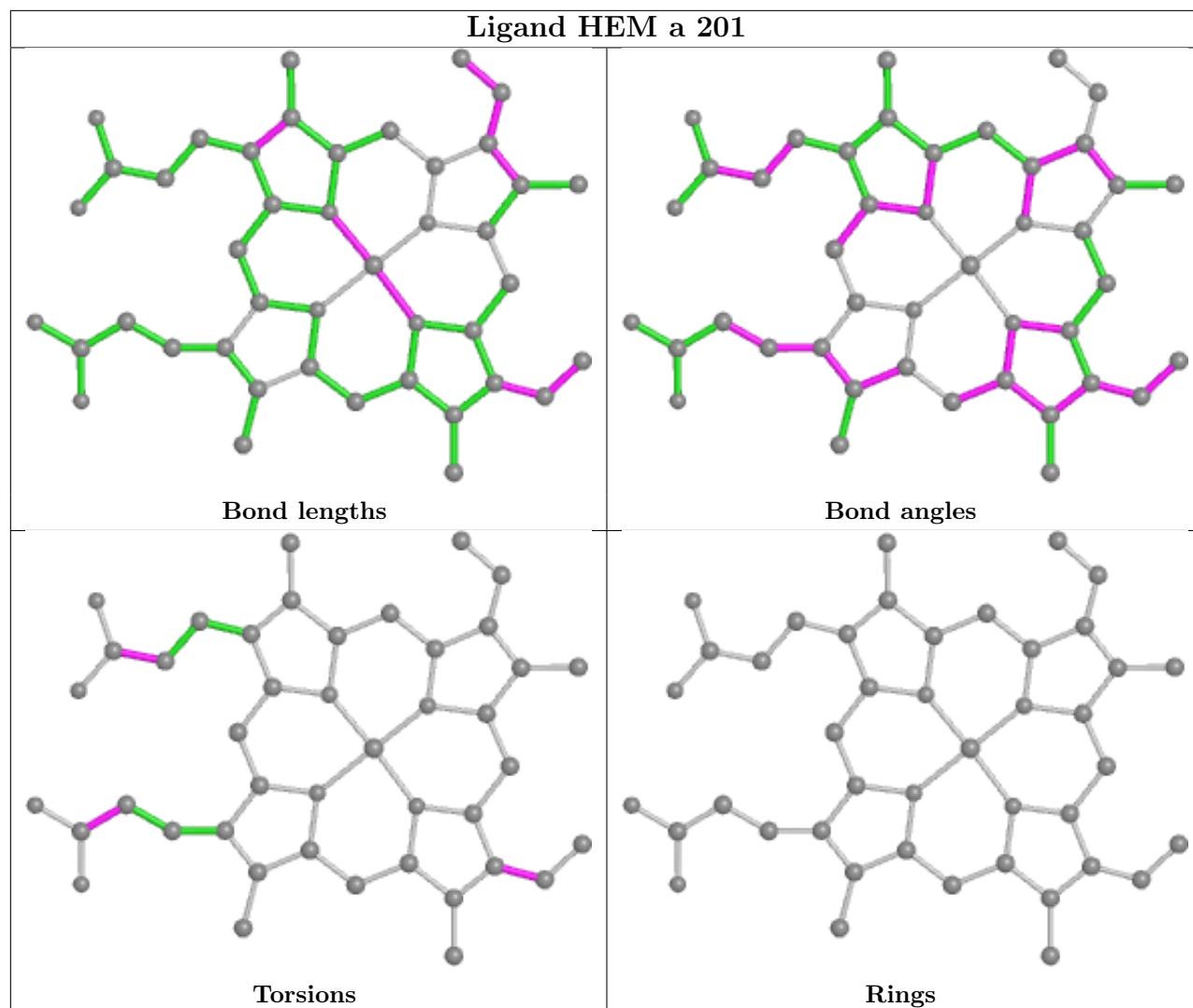












## 5.7 Other polymers [i](#)

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](#)

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	148/150 (98%)	0.01	2 (1%) 75 63	11, 32, 58, 80	0
1	E	148/150 (98%)	0.02	1 (0%) 87 81	7, 28, 55, 74	0
1	I	148/150 (98%)	-0.15	1 (0%) 87 81	6, 24, 54, 70	0
1	P	148/150 (98%)	0.11	5 (3%) 45 29	8, 34, 63, 78	0
1	T	148/150 (98%)	-0.10	1 (0%) 87 81	8, 25, 50, 68	0
1	X	148/150 (98%)	-0.27	0 100 100	6, 26, 55, 67	0
1	e	148/150 (98%)	0.16	4 (2%) 54 39	13, 40, 69, 84	0
1	i	148/150 (98%)	-0.15	0 100 100	8, 30, 55, 70	0
1	m	148/150 (98%)	-0.04	0 100 100	8, 25, 52, 73	0
2	B	142/142 (100%)	-0.05	3 (2%) 63 49	9, 23, 59, 88	0
2	F	142/142 (100%)	-0.27	1 (0%) 87 81	7, 22, 54, 106	0
2	J	142/142 (100%)	-0.07	3 (2%) 63 49	6, 24, 62, 98	0
2	Q	142/142 (100%)	-0.08	1 (0%) 87 81	7, 22, 64, 86	0
2	U	142/142 (100%)	-0.35	1 (0%) 87 81	5, 23, 53, 94	0
2	Y	142/142 (100%)	-0.06	2 (1%) 75 63	10, 26, 58, 95	0
2	f	142/142 (100%)	-0.18	1 (0%) 87 81	5, 25, 63, 89	0
2	j	142/142 (100%)	-0.30	2 (1%) 75 63	6, 24, 53, 101	0
2	n	142/142 (100%)	-0.08	3 (2%) 63 49	6, 26, 58, 93	0
3	C	151/151 (100%)	-0.19	1 (0%) 87 81	8, 27, 59, 92	0
3	G	151/151 (100%)	-0.08	4 (2%) 56 40	9, 28, 66, 96	0
3	K	151/151 (100%)	-0.09	1 (0%) 87 81	8, 36, 62, 85	0
3	R	151/151 (100%)	-0.09	0 100 100	5, 27, 64, 88	0
3	V	151/151 (100%)	0.22	6 (3%) 38 25	8, 30, 71, 92	0
3	Z	151/151 (100%)	-0.13	2 (1%) 77 65	9, 35, 65, 83	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
3	g	151/151 (100%)	-0.17	0 100 100	7, 28, 57, 88	0
3	k	151/151 (100%)	0.12	5 (3%) 46 30	11, 35, 72, 98	0
3	o	151/151 (100%)	0.01	3 (1%) 65 51	10, 37, 70, 82	0
4	D	141/141 (100%)	-0.07	2 (1%) 75 63	6, 30, 61, 80	0
4	H	141/141 (100%)	-0.25	1 (0%) 87 81	10, 24, 57, 81	0
4	L	141/141 (100%)	-0.18	1 (0%) 87 81	8, 26, 55, 83	0
4	S	141/141 (100%)	-0.05	1 (0%) 87 81	13, 32, 62, 79	0
4	W	141/141 (100%)	-0.24	1 (0%) 87 81	8, 30, 55, 79	0
4	a	141/141 (100%)	0.13	0 100 100	10, 33, 57, 82	0
4	h	141/141 (100%)	-0.13	1 (0%) 87 81	10, 35, 69, 89	0
4	l	141/141 (100%)	-0.16	0 100 100	10, 31, 61, 84	0
4	p	141/141 (100%)	-0.13	0 100 100	13, 29, 63, 77	0
5	M	222/224 (99%)	0.13	1 (0%) 91 86	6, 25, 65, 88	0
5	b	222/224 (99%)	0.10	1 (0%) 91 86	7, 25, 67, 89	0
5	q	222/224 (99%)	0.06	2 (0%) 84 75	7, 24, 70, 84	0
6	N	219/236 (92%)	0.07	1 (0%) 91 86	5, 23, 59, 88	0
6	c	219/236 (92%)	0.10	1 (0%) 91 86	5, 24, 56, 87	0
6	r	219/236 (92%)	0.01	1 (0%) 91 86	4, 24, 59, 90	0
7	O	213/218 (97%)	-0.01	1 (0%) 91 86	6, 24, 49, 84	0
7	d	213/218 (97%)	0.07	0 100 100	4, 25, 55, 89	0
7	s	213/218 (97%)	-0.08	1 (0%) 91 86	6, 23, 55, 85	0
All	All	7200/7290 (98%)	-0.05	69 (0%) 82 72	4, 27, 62, 106	0

The worst 5 of 69 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	V	1	HIS	6.8
3	G	1	HIS	5.8
3	V	2	GLN	5.5
1	e	150	HIS	5.3
2	Q	1	SER	4.6

## 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(Å <sup>2</sup> )	Q<0.9
11	ZN	d	302	1/1	0.80	0.25	28,28,28,28	0
12	NAG	s	303	14/15	0.86	0.30	28,28,28,28	0
12	NAG	O	304	14/15	0.88	0.34	28,28,28,28	0
12	NAG	d	304	14/15	0.89	0.40	28,28,28,28	0
11	ZN	s	302	1/1	0.89	0.25	28,28,28,28	0
8	HEM	e	201	43/43	0.90	0.27	28,28,28,28	0
11	ZN	O	302	1/1	0.91	0.30	28,28,28,28	0
8	HEM	V	201	43/43	0.94	0.28	28,28,28,28	0
8	HEM	T	201	43/43	0.95	0.25	28,28,28,28	0
8	HEM	i	201	43/43	0.95	0.23	28,28,28,28	0
8	HEM	k	201	43/43	0.95	0.29	28,28,28,28	0
8	HEM	n	201	43/43	0.95	0.26	28,28,28,28	0
8	HEM	J	201	43/43	0.95	0.25	28,28,28,28	0
8	HEM	a	201	43/43	0.96	0.31	28,28,28,28	0
8	HEM	L	201	43/43	0.96	0.25	28,28,28,28	0
8	HEM	f	201	43/43	0.96	0.24	28,28,28,28	0
8	HEM	g	201	43/43	0.96	0.25	28,28,28,28	0
8	HEM	P	201	43/43	0.96	0.22	28,28,28,28	0
8	HEM	Q	201	43/43	0.96	0.23	28,28,28,28	0
8	HEM	l	201	43/43	0.96	0.25	28,28,28,28	0
8	HEM	m	201	43/43	0.96	0.26	28,28,28,28	0
8	HEM	R	201	43/43	0.96	0.26	28,28,28,28	0
8	HEM	p	201	43/43	0.96	0.25	28,28,28,28	0
9	CYN	C	202	2/2	0.96	0.32	28,28,28,28	0
8	HEM	S	201	43/43	0.96	0.26	28,28,28,28	0
8	HEM	E	201	43/43	0.96	0.30	28,28,28,28	0
8	HEM	K	201	43/43	0.96	0.21	28,28,28,28	0

*Continued on next page...*



*Continued from previous page...*

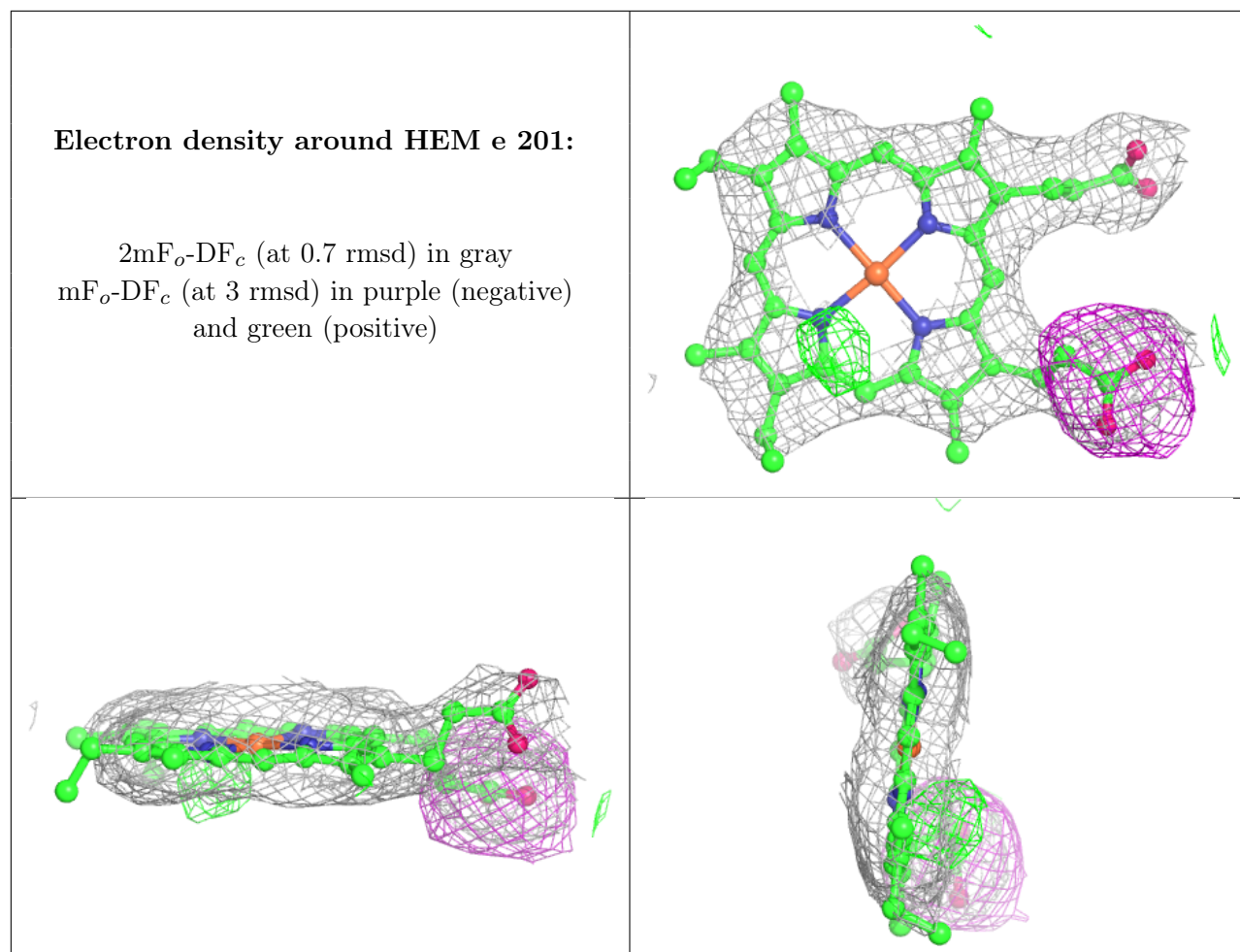
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
8	HEM	X	201	43/43	0.96	0.24	28,28,28,28	0
8	HEM	Y	201	43/43	0.96	0.26	28,28,28,28	0
8	HEM	Z	201	43/43	0.96	0.22	28,28,28,28	0
8	HEM	B	201	43/43	0.97	0.26	28,28,28,28	0
8	HEM	W	201	43/43	0.97	0.23	28,28,28,28	0
8	HEM	C	201	43/43	0.97	0.23	28,28,28,28	0
8	HEM	D	201	43/43	0.97	0.26	28,28,28,28	0
8	HEM	o	201	43/43	0.97	0.23	28,28,28,28	0
8	HEM	A	201	43/43	0.97	0.27	28,28,28,28	0
8	HEM	F	201	43/43	0.97	0.23	28,28,28,28	0
9	CYN	R	202	2/2	0.97	0.32	28,28,28,28	0
9	CYN	W	202	2/2	0.97	0.21	28,28,28,28	0
10	CA	d	301	1/1	0.97	0.10	28,28,28,28	0
8	HEM	G	201	43/43	0.97	0.21	28,28,28,28	0
8	HEM	H	201	43/43	0.97	0.21	28,28,28,28	0
8	HEM	I	201	43/43	0.97	0.28	28,28,28,28	0
8	HEM	h	201	43/43	0.97	0.26	28,28,28,28	0
8	HEM	U	201	43/43	0.97	0.23	28,28,28,28	0
8	HEM	j	201	43/43	0.97	0.24	28,28,28,28	0
9	CYN	Z	202	2/2	0.98	0.18	28,28,28,28	0
9	CYN	i	202	2/2	0.98	0.17	28,28,28,28	0
9	CYN	j	202	2/2	0.98	0.19	28,28,28,28	0
9	CYN	l	202	2/2	0.98	0.23	28,28,28,28	0
10	CA	M	301	1/1	0.98	0.08	28,28,28,28	0
10	CA	N	301	1/1	0.98	0.11	28,28,28,28	0
10	CA	O	301	1/1	0.98	0.06	28,28,28,28	0
10	CA	b	301	1/1	0.98	0.08	28,28,28,28	0
9	CYN	G	202	2/2	0.98	0.18	28,28,28,28	0
10	CA	q	301	1/1	0.98	0.07	28,28,28,28	0
9	CYN	P	202	2/2	0.98	0.25	28,28,28,28	0
9	CYN	E	202	2/2	0.98	0.18	28,28,28,28	0
11	ZN	d	303	1/1	0.98	0.05	28,28,28,28	0
9	CYN	T	202	2/2	0.98	0.18	28,28,28,28	0
9	CYN	U	202	2/2	0.98	0.17	28,28,28,28	0
9	CYN	F	202	2/2	0.98	0.20	28,28,28,28	0
9	CYN	Y	202	2/2	0.98	0.21	28,28,28,28	0
9	CYN	L	202	2/2	0.99	0.21	28,28,28,28	0
9	CYN	B	202	2/2	0.99	0.17	28,28,28,28	0
9	CYN	k	202	2/2	0.99	0.10	28,28,28,28	0
9	CYN	A	202	2/2	0.99	0.22	28,28,28,28	0
9	CYN	m	202	2/2	0.99	0.18	28,28,28,28	0
9	CYN	n	202	2/2	0.99	0.16	28,28,28,28	0

*Continued on next page...*

*Continued from previous page...*

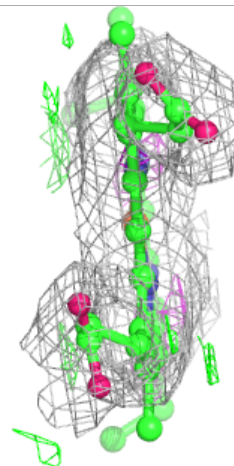
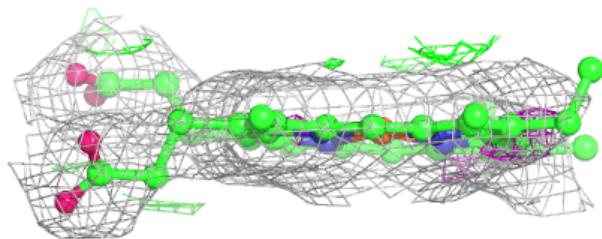
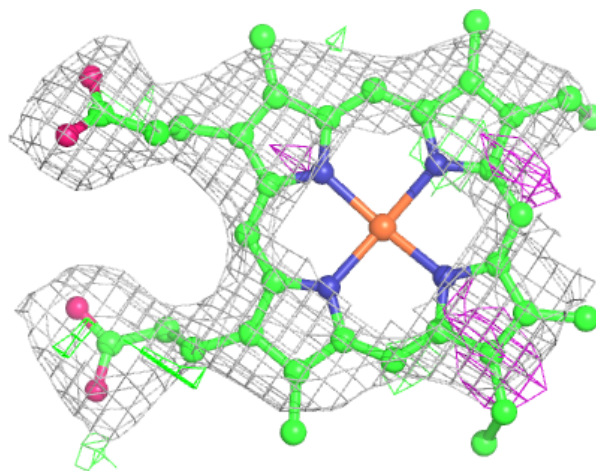
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
9	CYN	o	202	2/2	0.99	0.11	28,28,28,28	0
9	CYN	p	202	2/2	0.99	0.15	28,28,28,28	0
9	CYN	S	202	2/2	0.99	0.26	28,28,28,28	0
9	CYN	D	202	2/2	0.99	0.18	28,28,28,28	0
9	CYN	H	202	2/2	0.99	0.20	28,28,28,28	0
9	CYN	V	202	2/2	0.99	0.15	28,28,28,28	0
10	CA	c	301	1/1	0.99	0.08	28,28,28,28	0
9	CYN	I	202	2/2	0.99	0.19	28,28,28,28	0
9	CYN	X	202	2/2	0.99	0.23	28,28,28,28	0
10	CA	r	301	1/1	0.99	0.09	28,28,28,28	0
10	CA	s	301	1/1	0.99	0.09	28,28,28,28	0
9	CYN	J	202	2/2	0.99	0.18	28,28,28,28	0
11	ZN	O	303	1/1	0.99	0.09	28,28,28,28	0
11	ZN	c	302	1/1	0.99	0.07	28,28,28,28	0
9	CYN	K	202	2/2	0.99	0.17	28,28,28,28	0
9	CYN	a	202	2/2	0.99	0.14	28,28,28,28	0
9	CYN	e	202	2/2	0.99	0.18	28,28,28,28	0
9	CYN	f	202	2/2	0.99	0.18	28,28,28,28	0
9	CYN	g	202	2/2	0.99	0.26	28,28,28,28	0
9	CYN	h	202	2/2	0.99	0.30	28,28,28,28	0
9	CYN	Q	202	2/2	1.00	0.12	28,28,28,28	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.



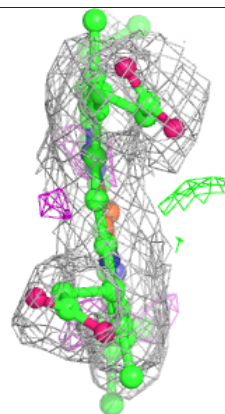
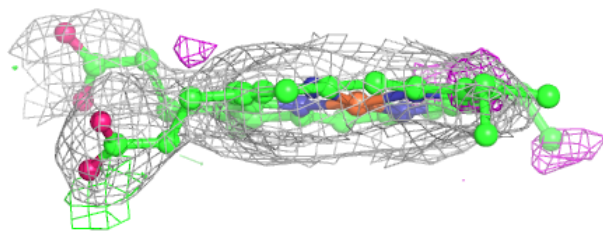
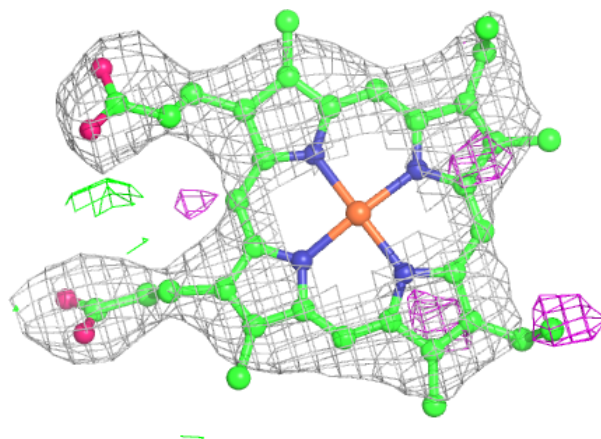
**Electron density around HEM V 201:**

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

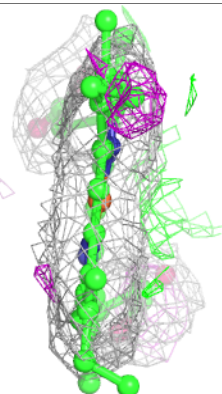
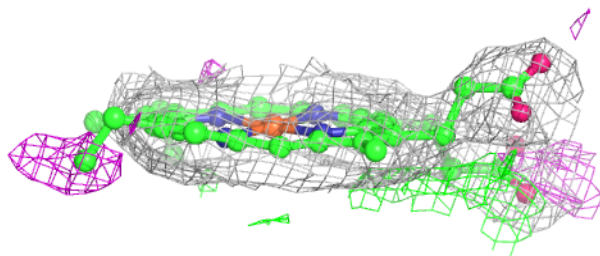
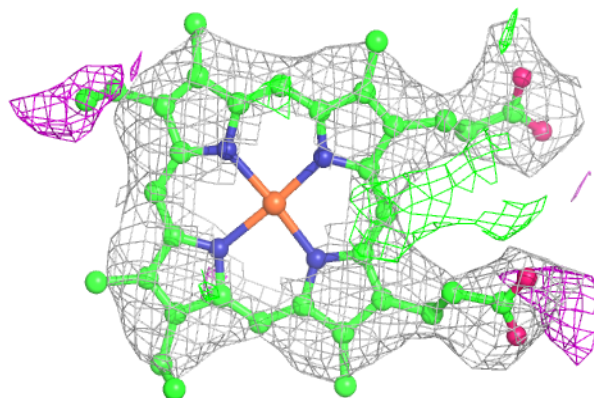


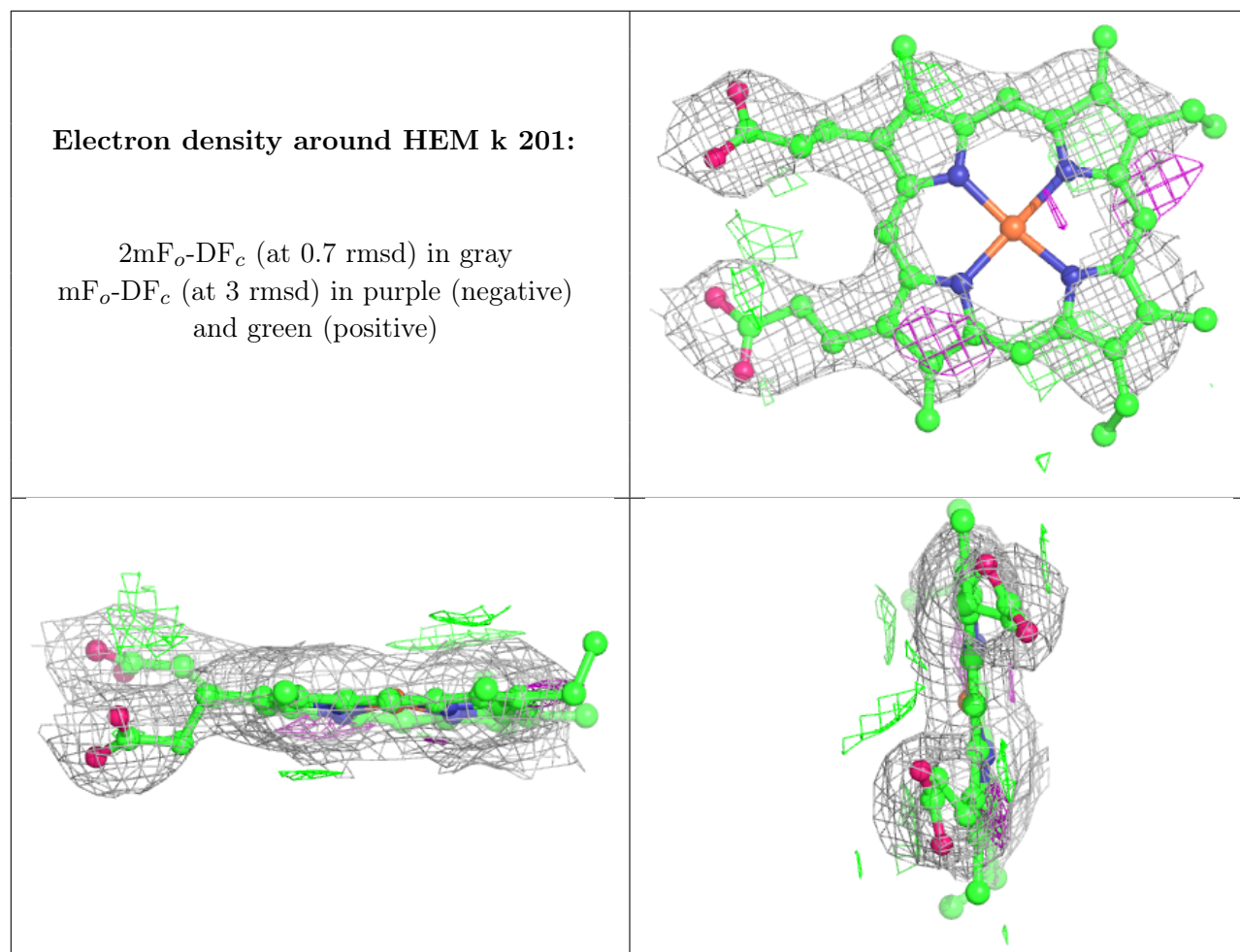
**Electron density around HEM T 201:**

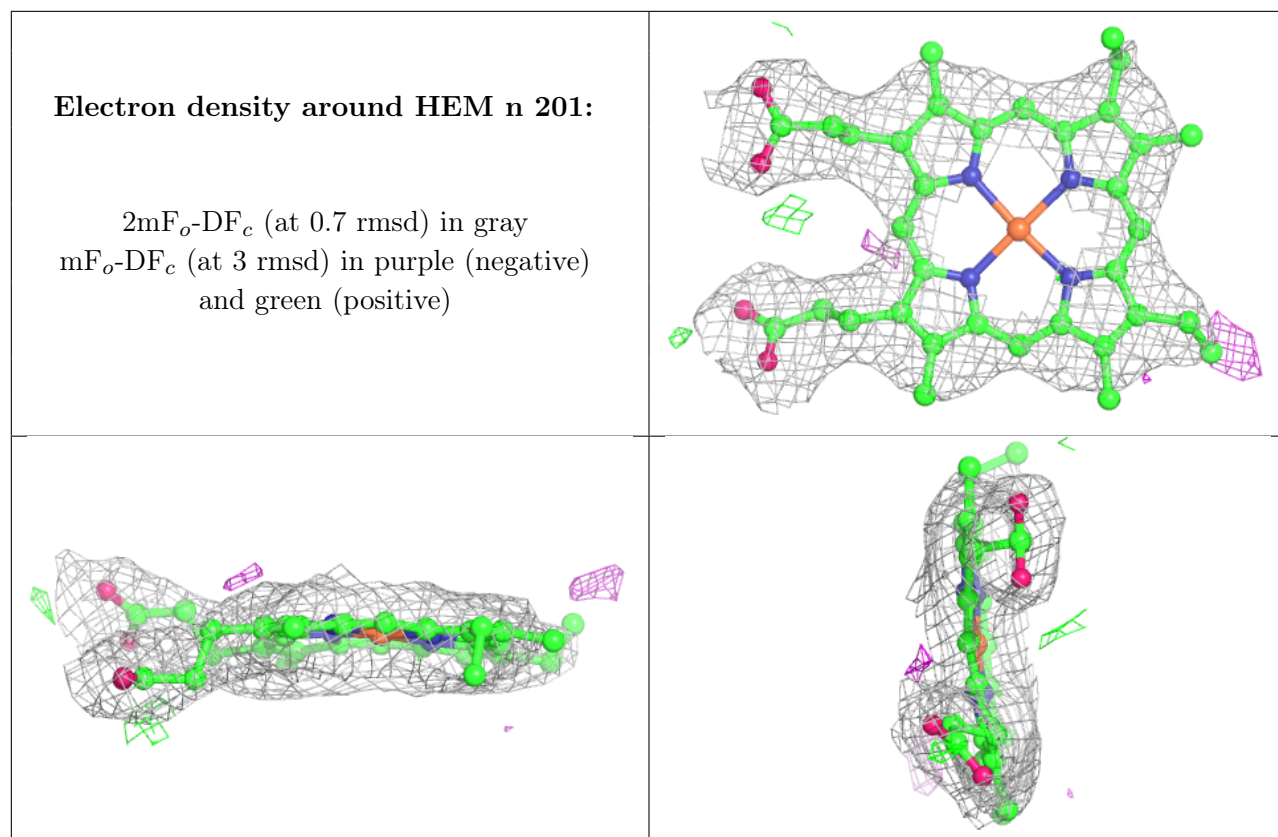
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

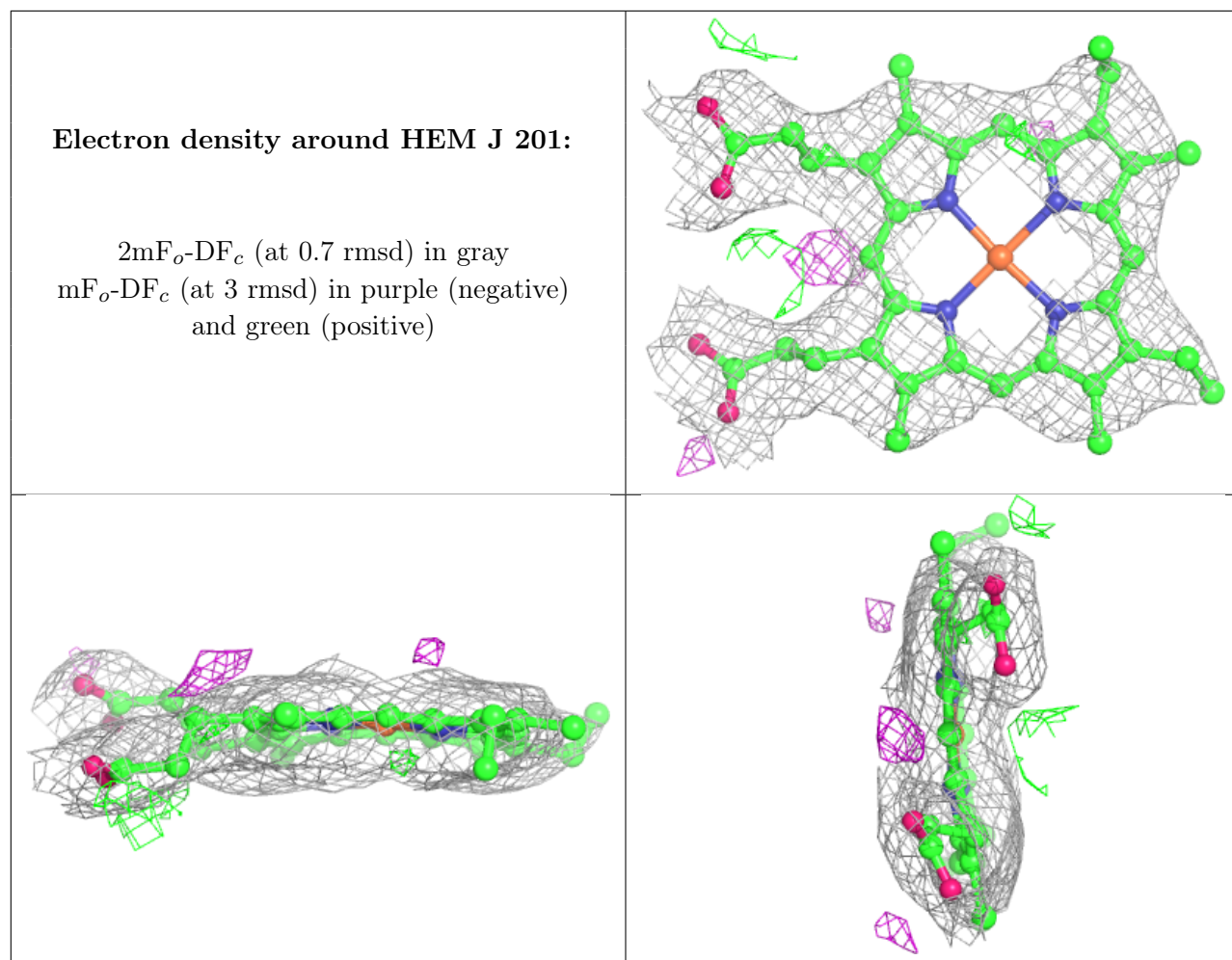
**Electron density around HEM i 201:**

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

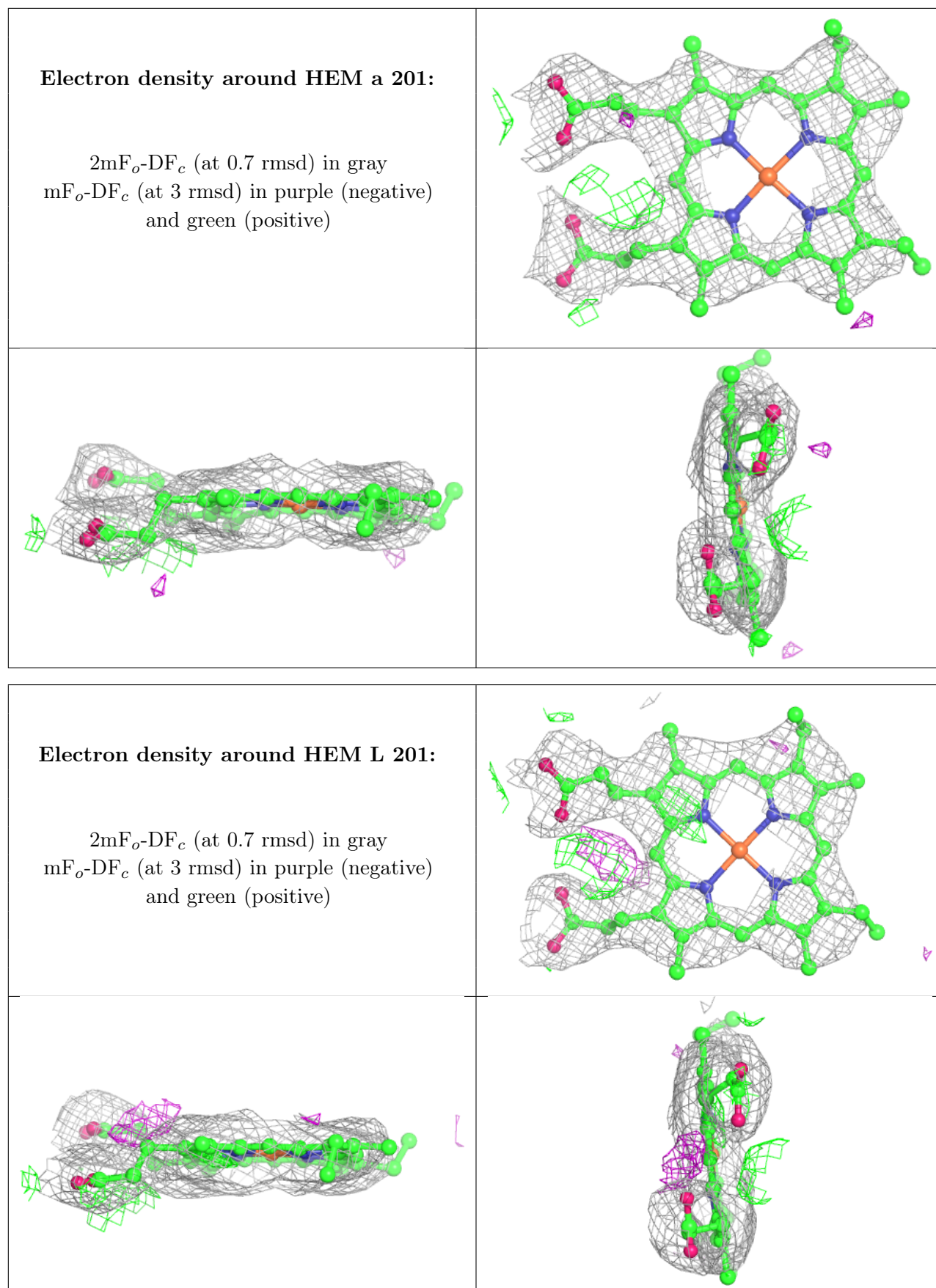






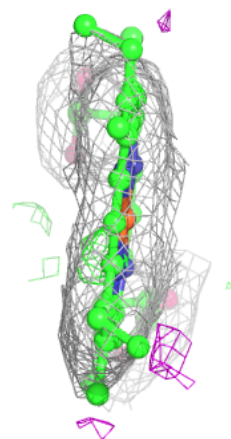
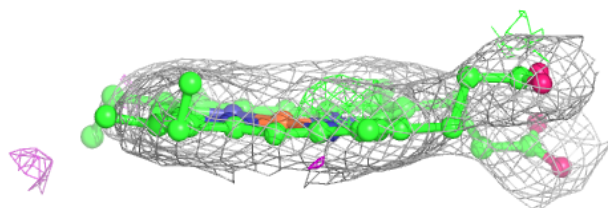
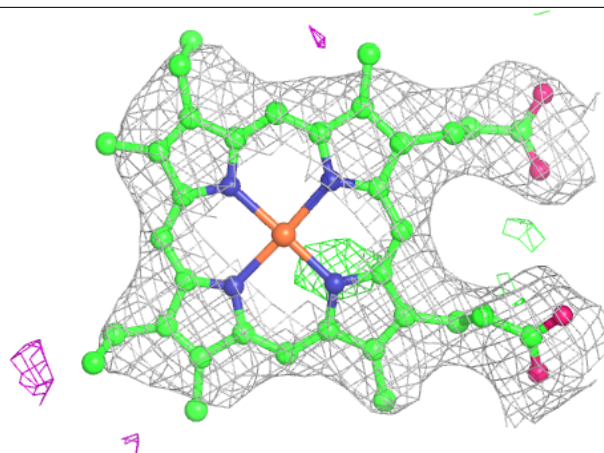


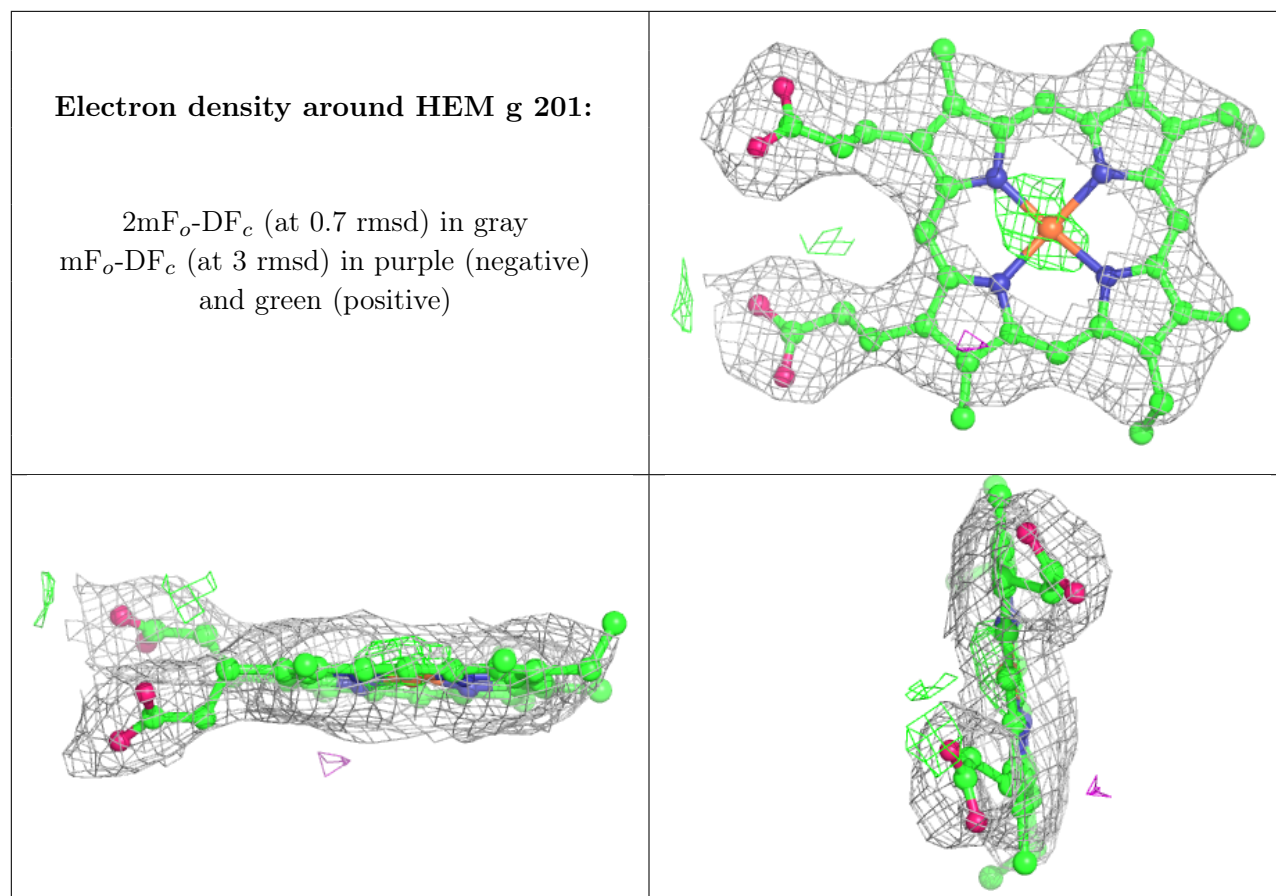




**Electron density around HEM f 201:**

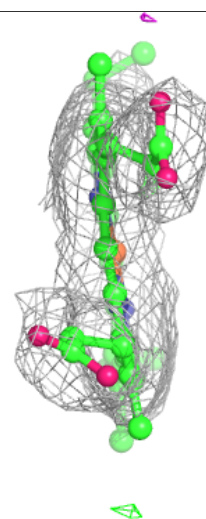
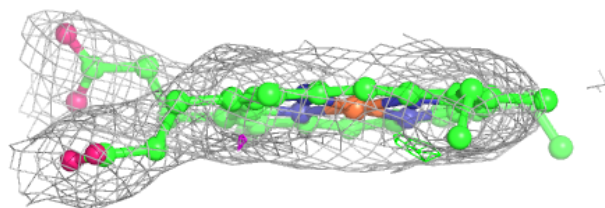
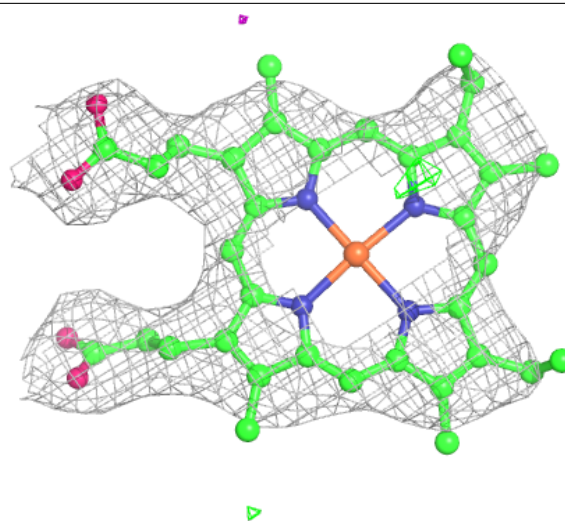
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





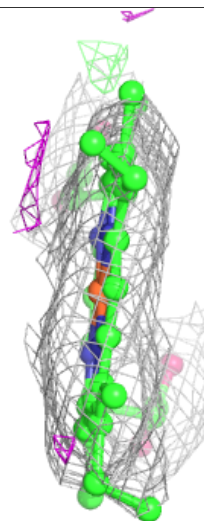
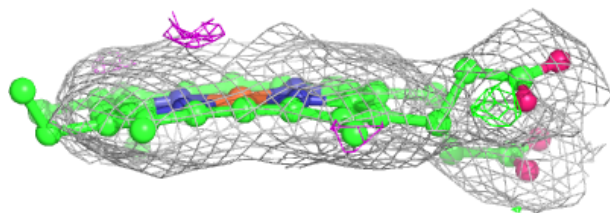
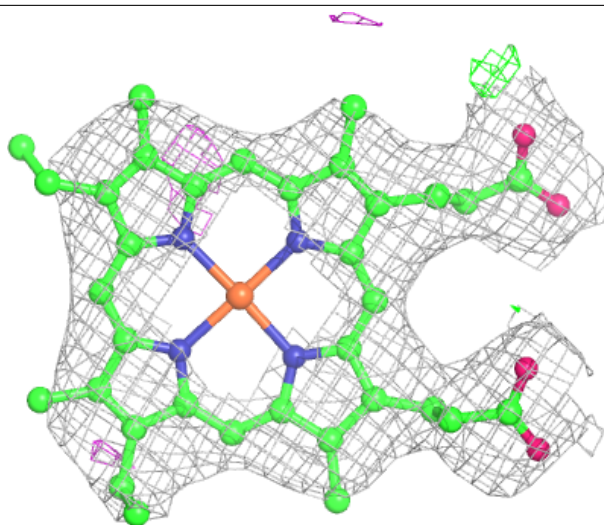
**Electron density around HEM P 201:**

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



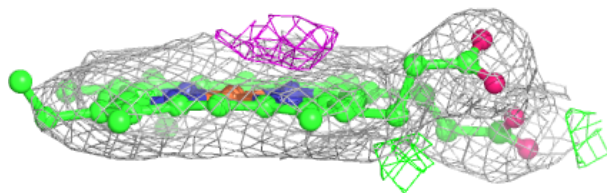
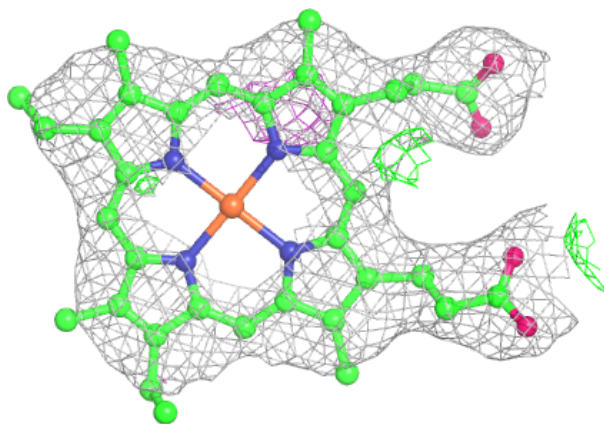
**Electron density around HEM Q 201:**

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



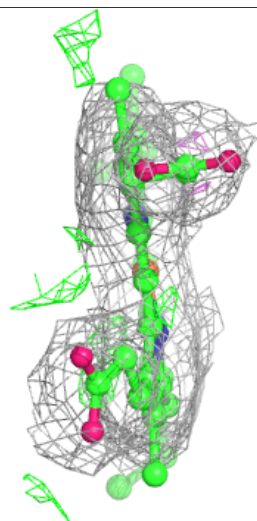
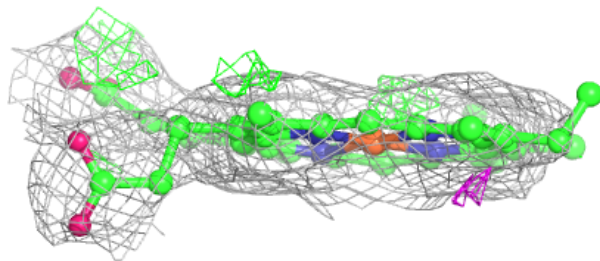
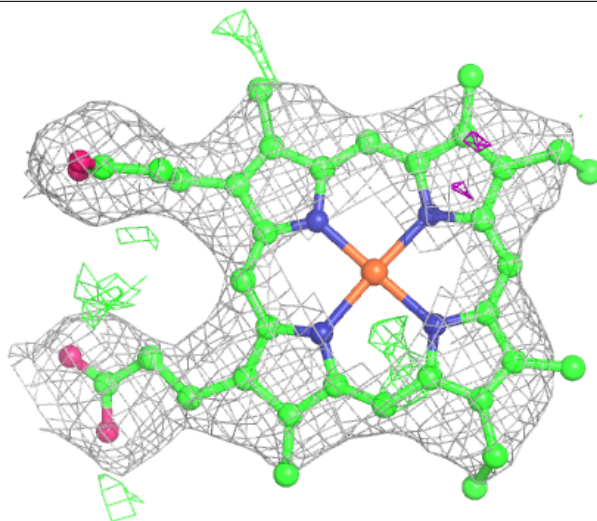
**Electron density around HEM 1 201:**

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



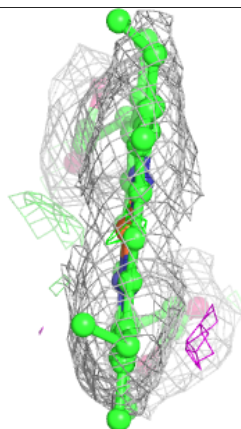
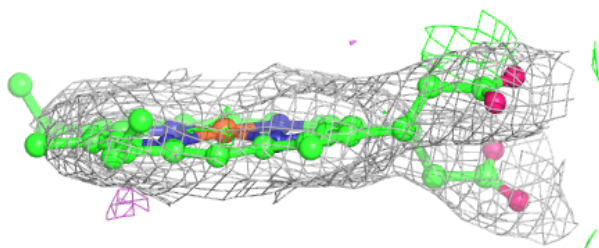
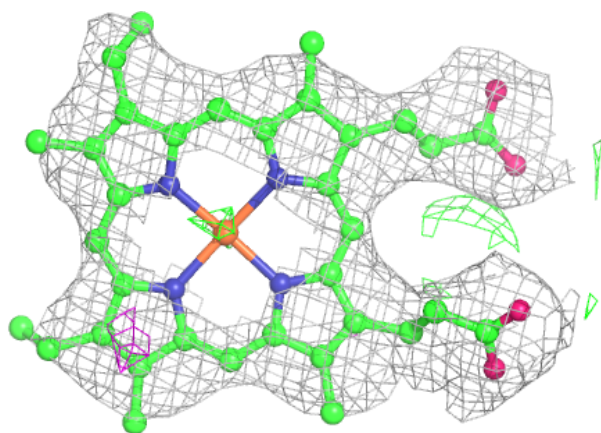
**Electron density around HEM m 201:**

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

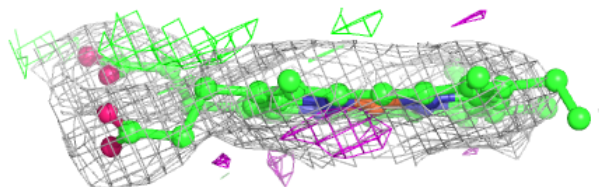
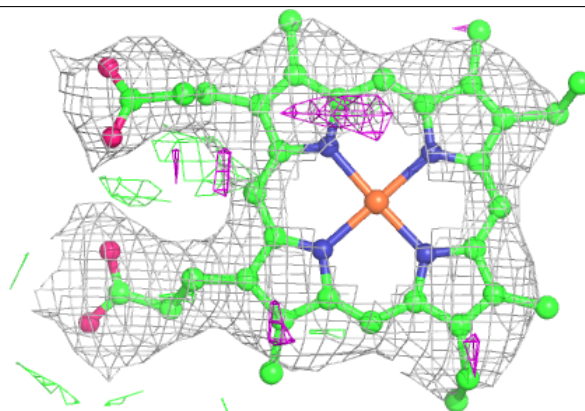


**Electron density around HEM R 201:**

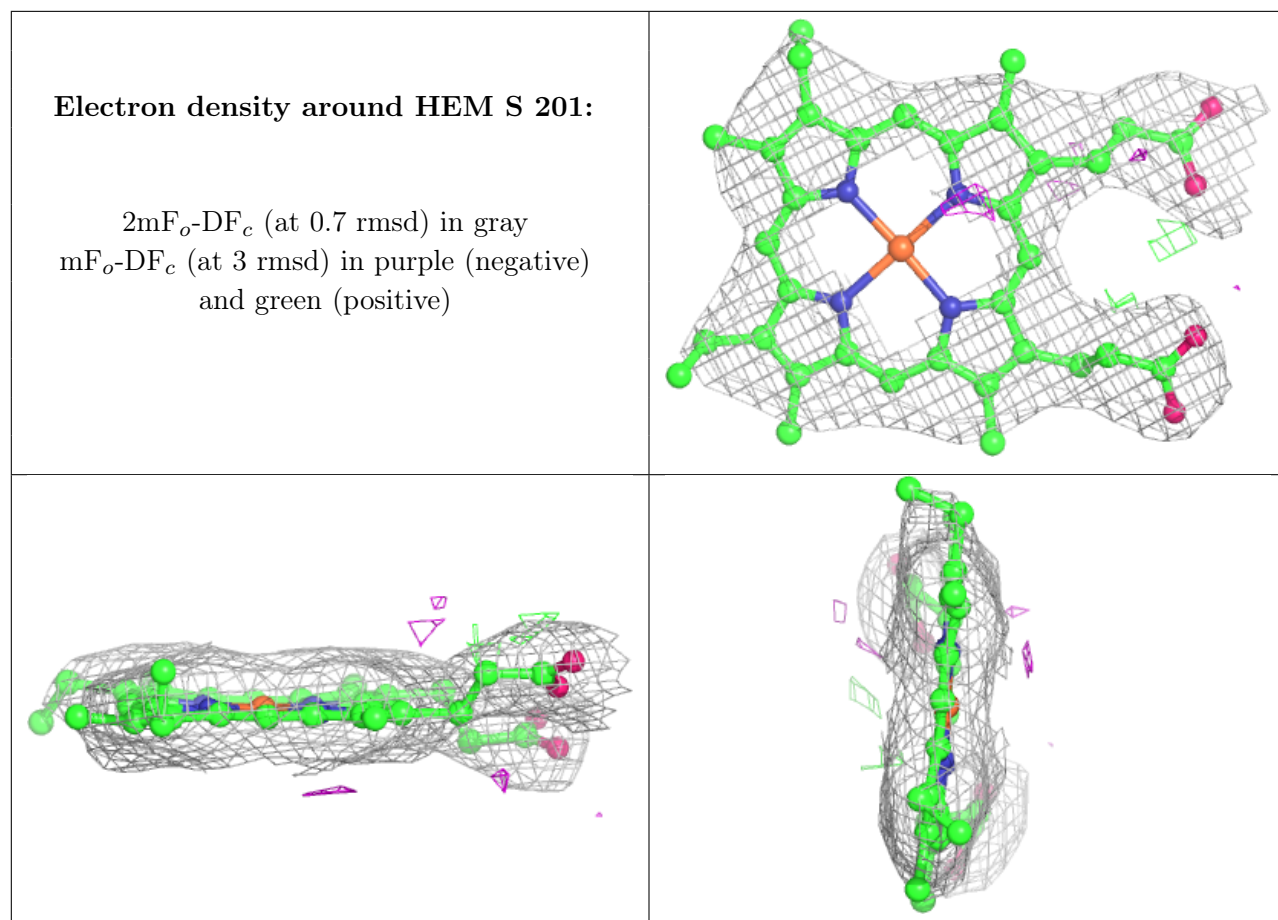
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around HEM p 201:**

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

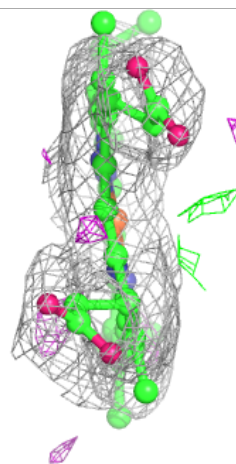
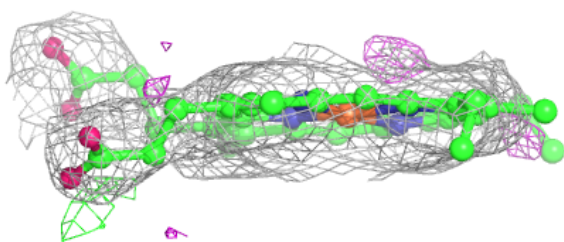
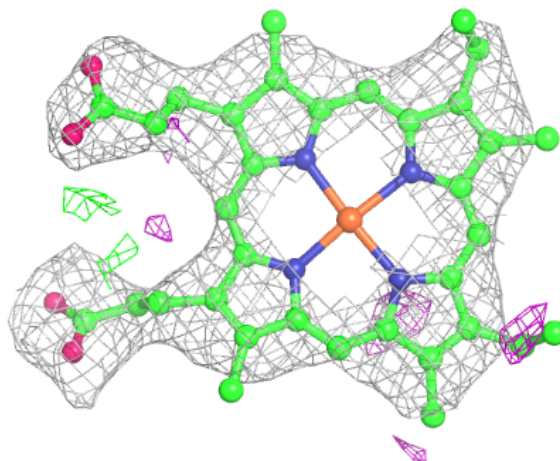


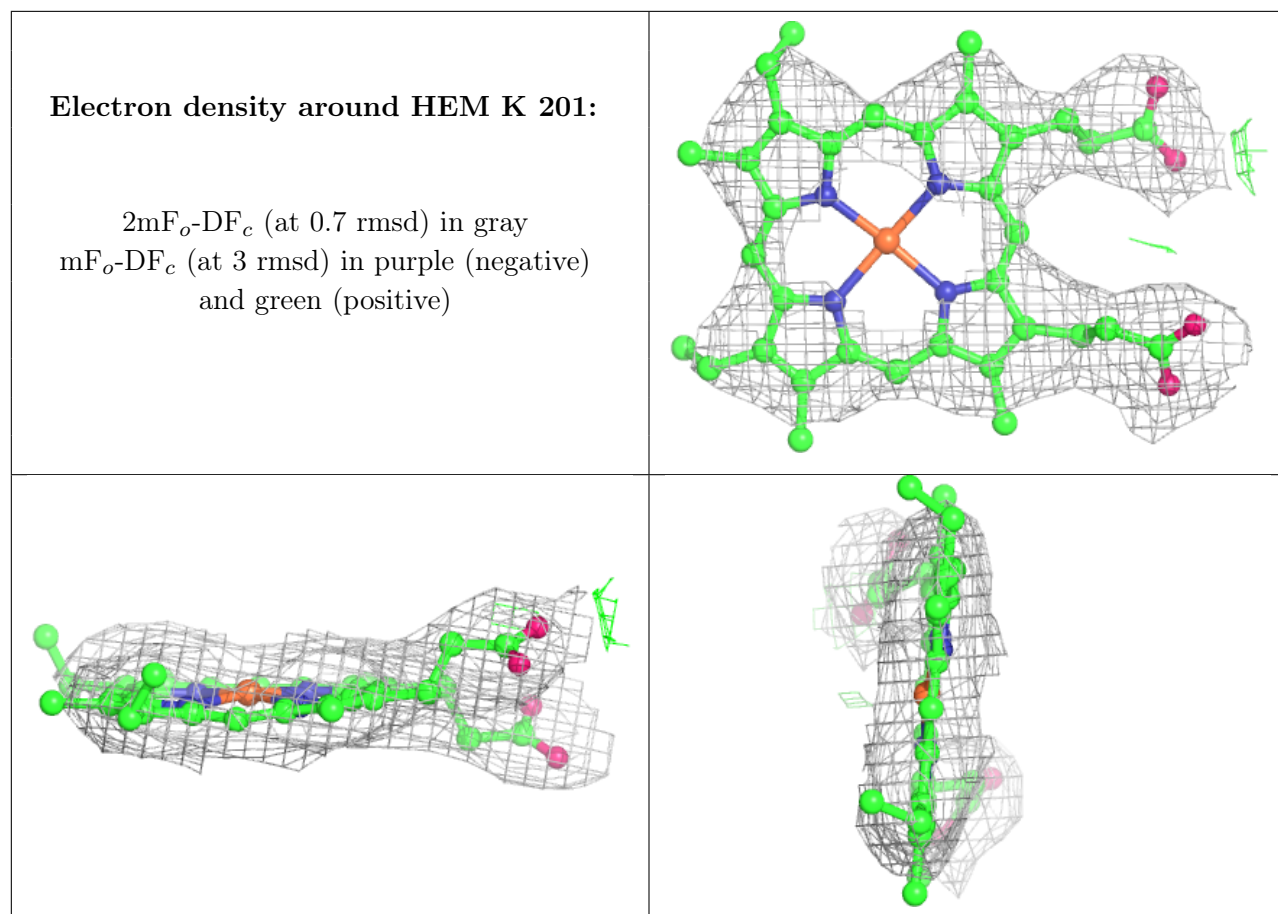


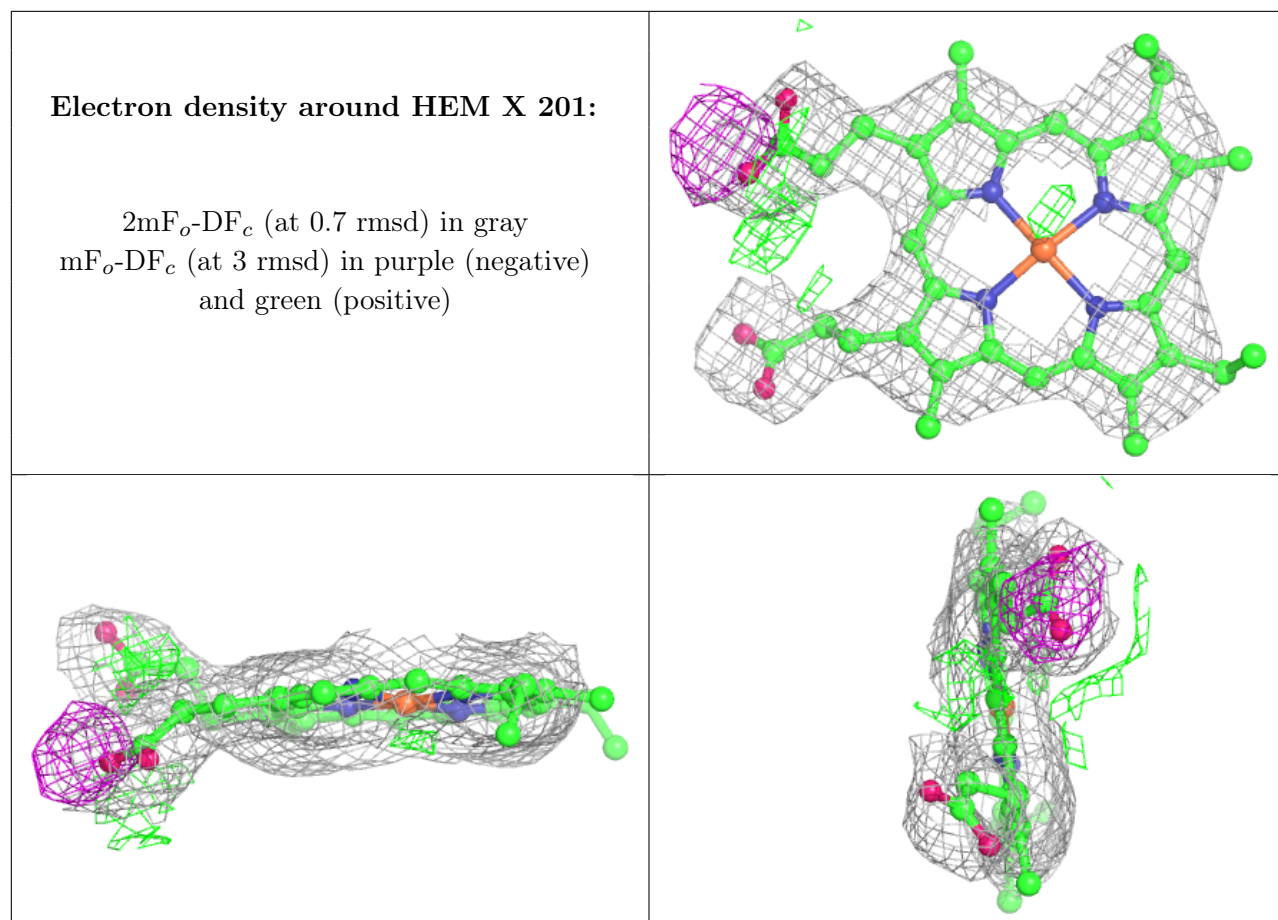


**Electron density around HEM E 201:**

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

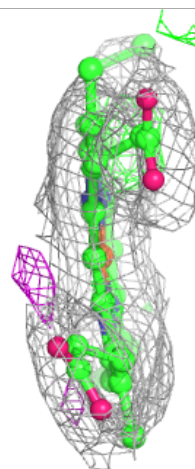
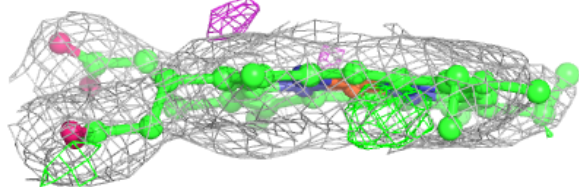
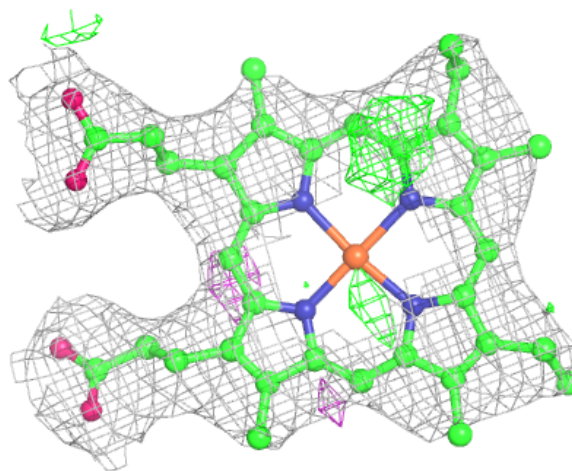






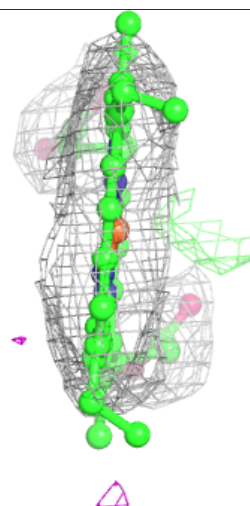
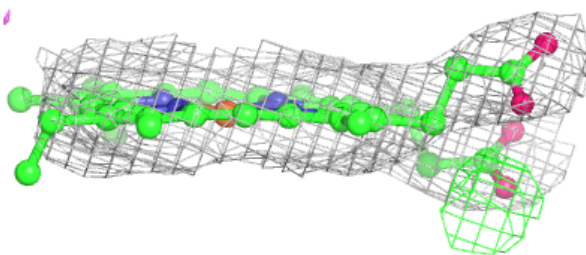
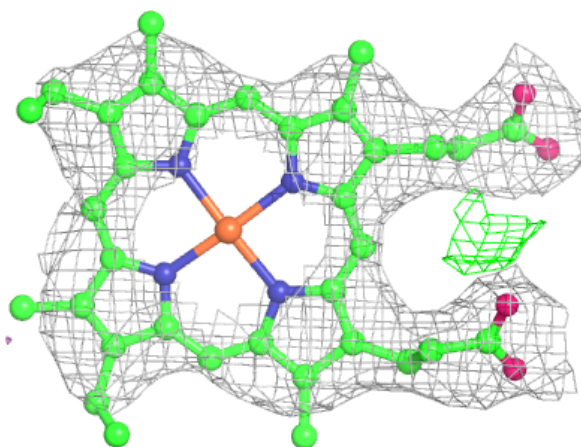
**Electron density around HEM Y 201:**

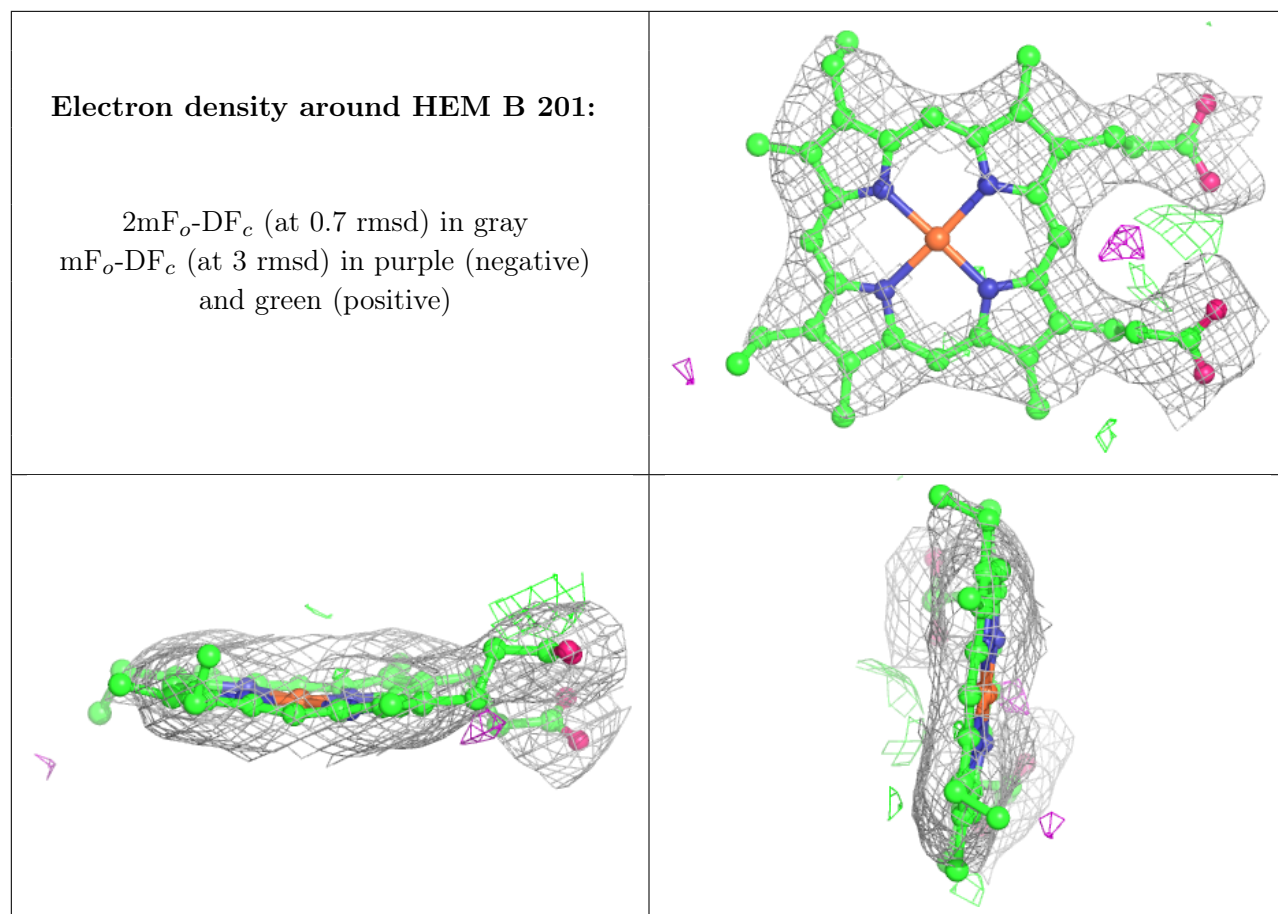
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

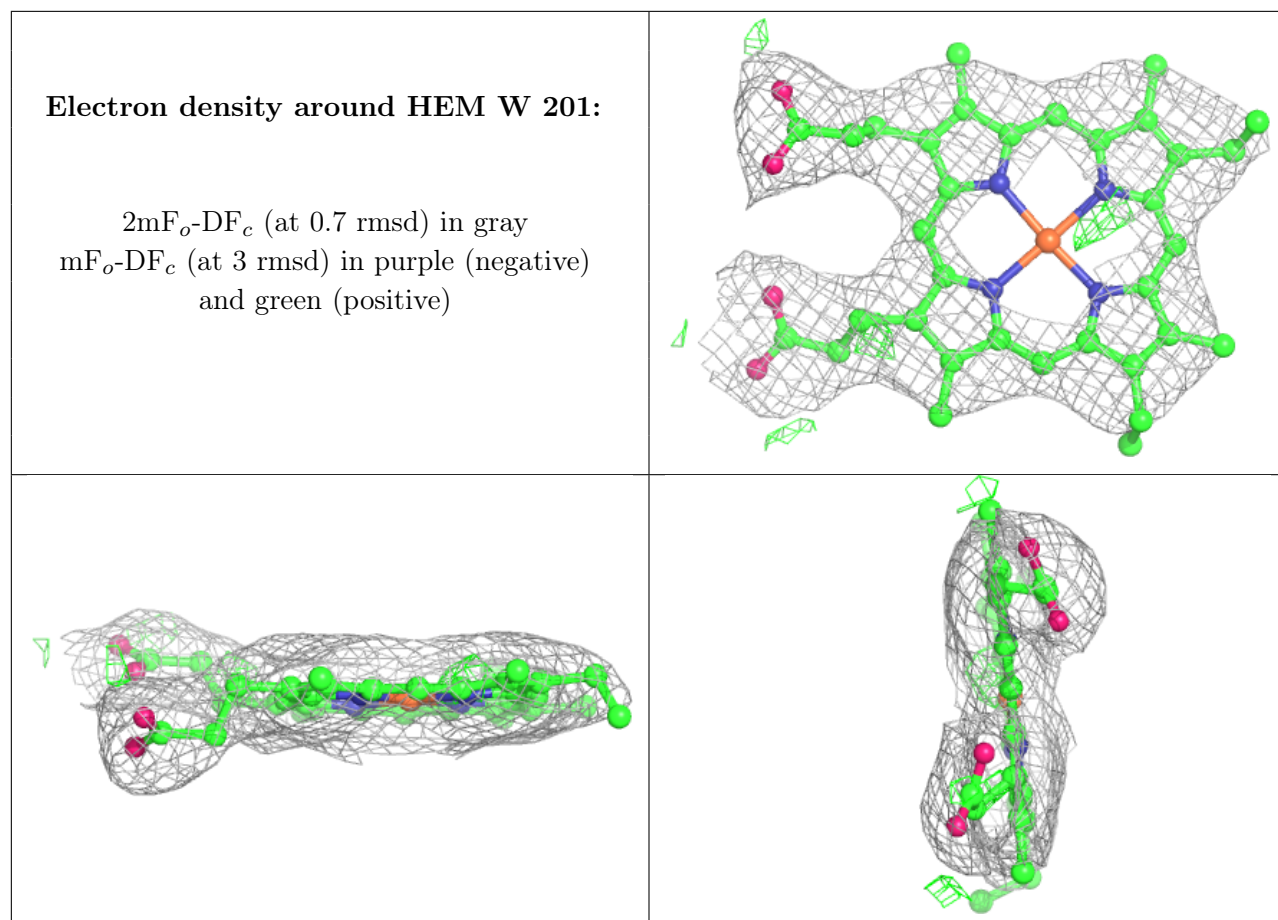


**Electron density around HEM Z 201:**

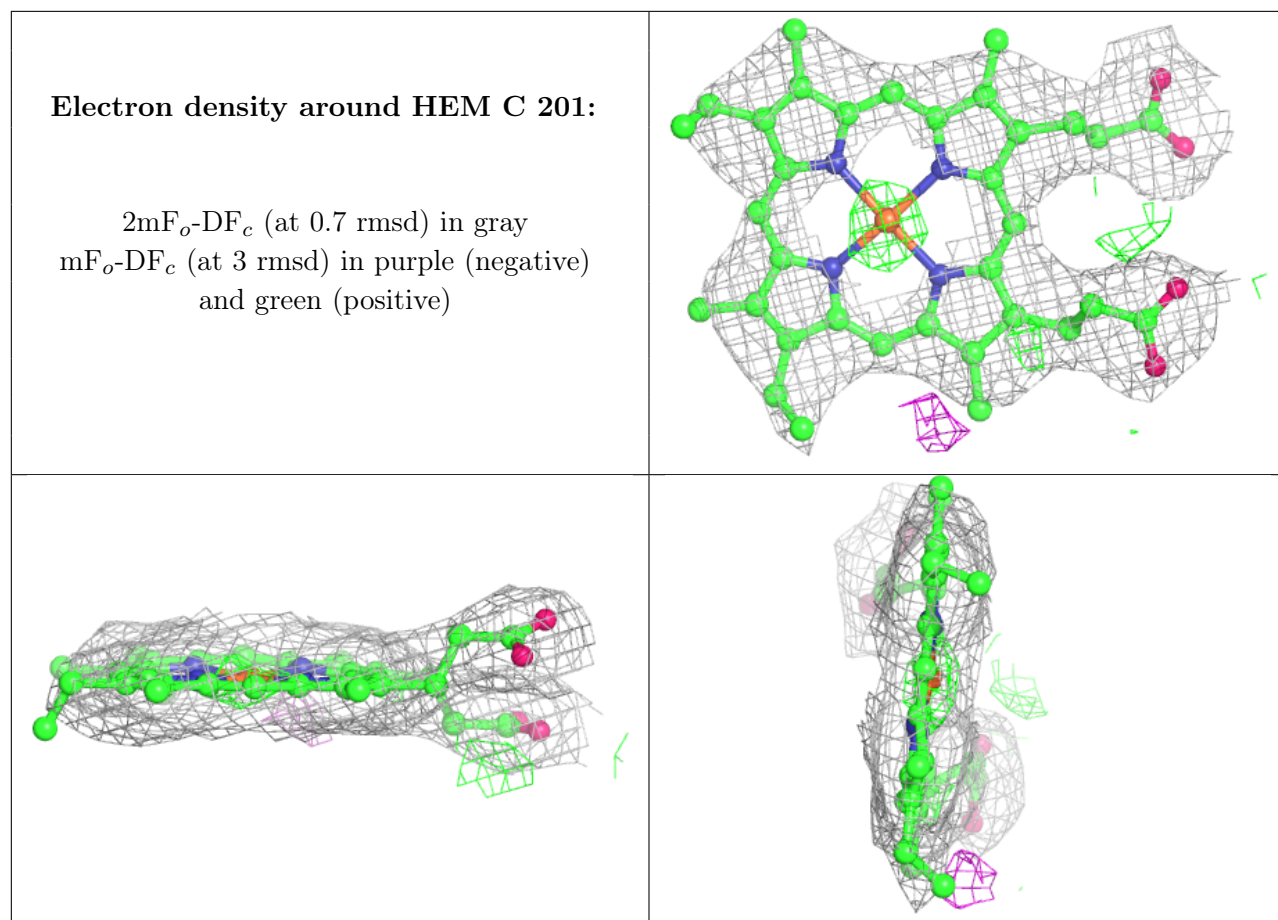
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





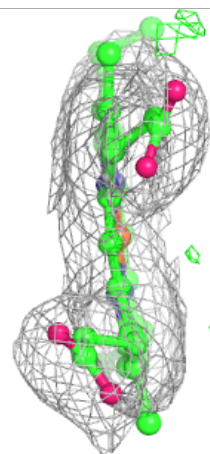
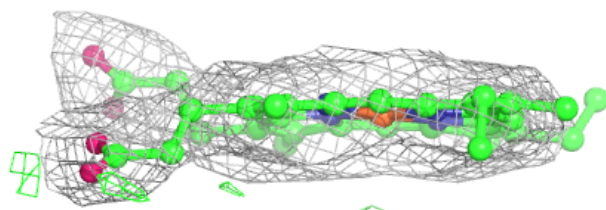
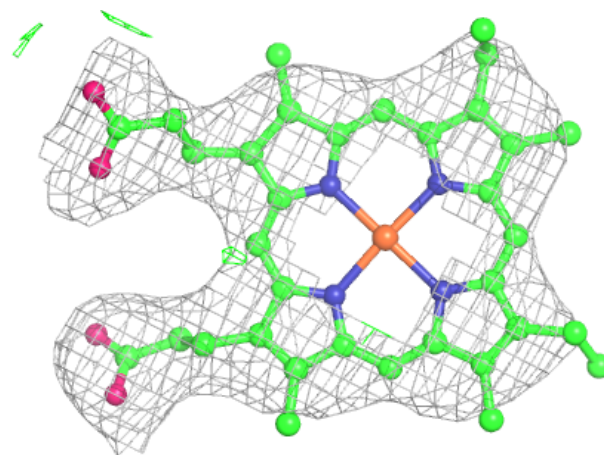


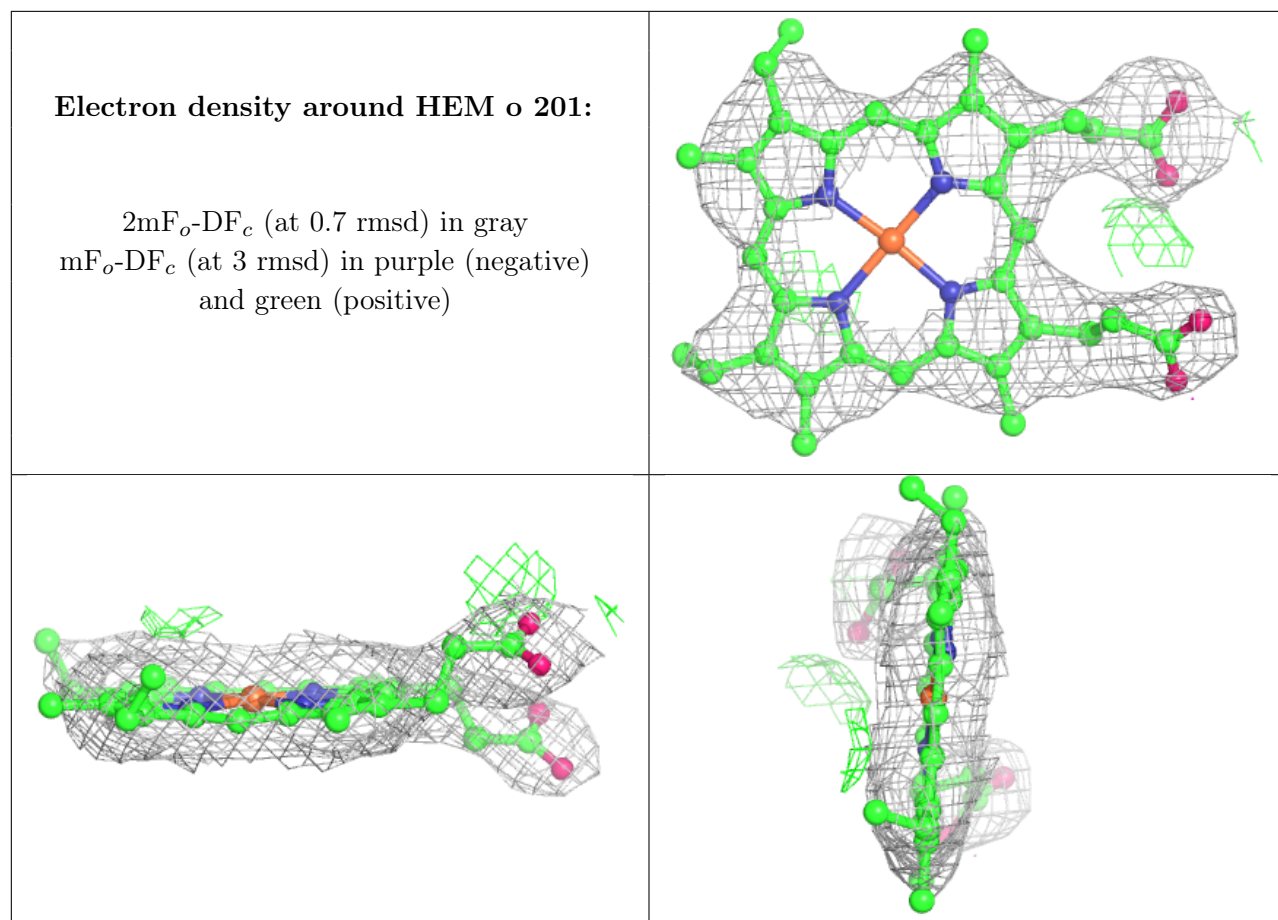


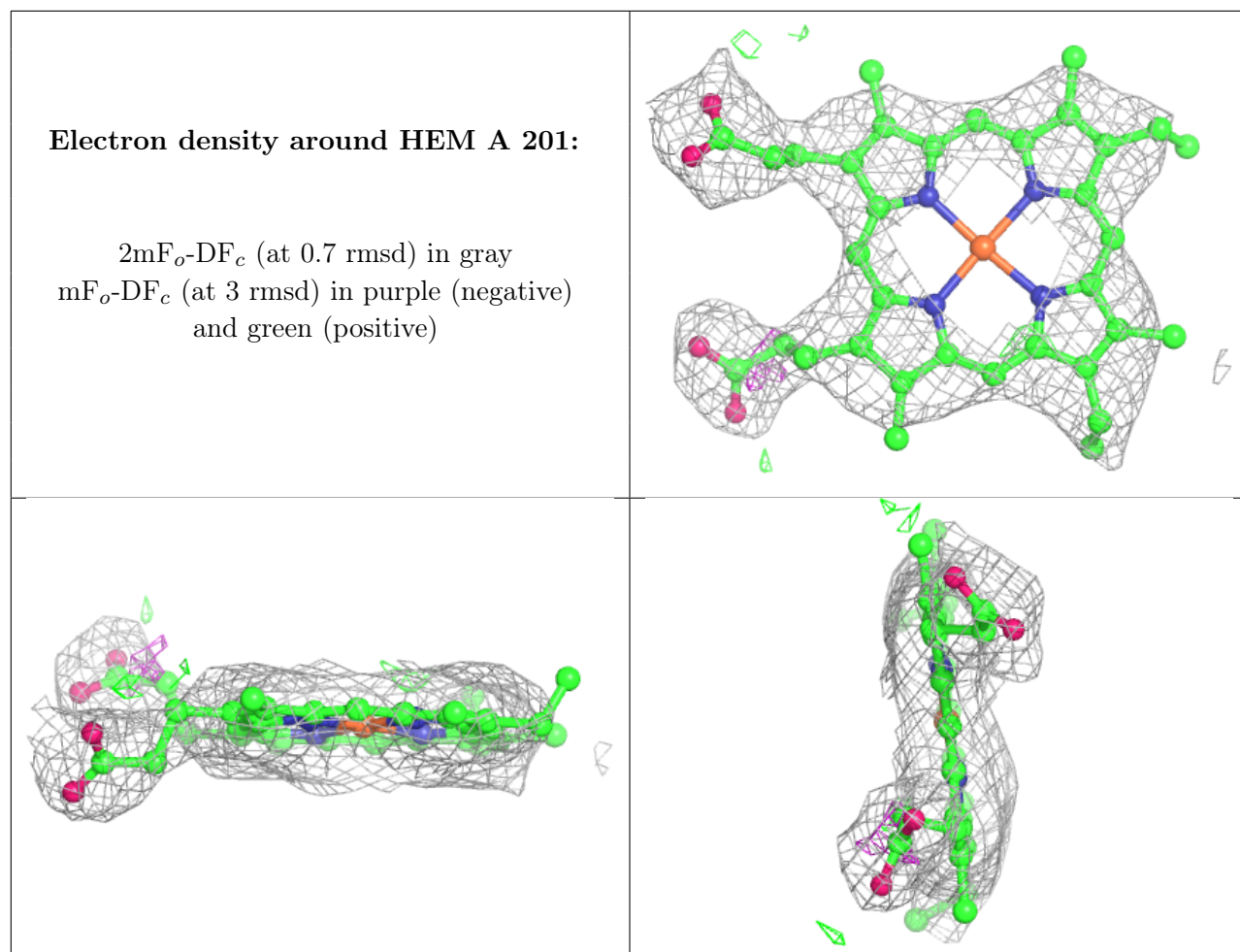


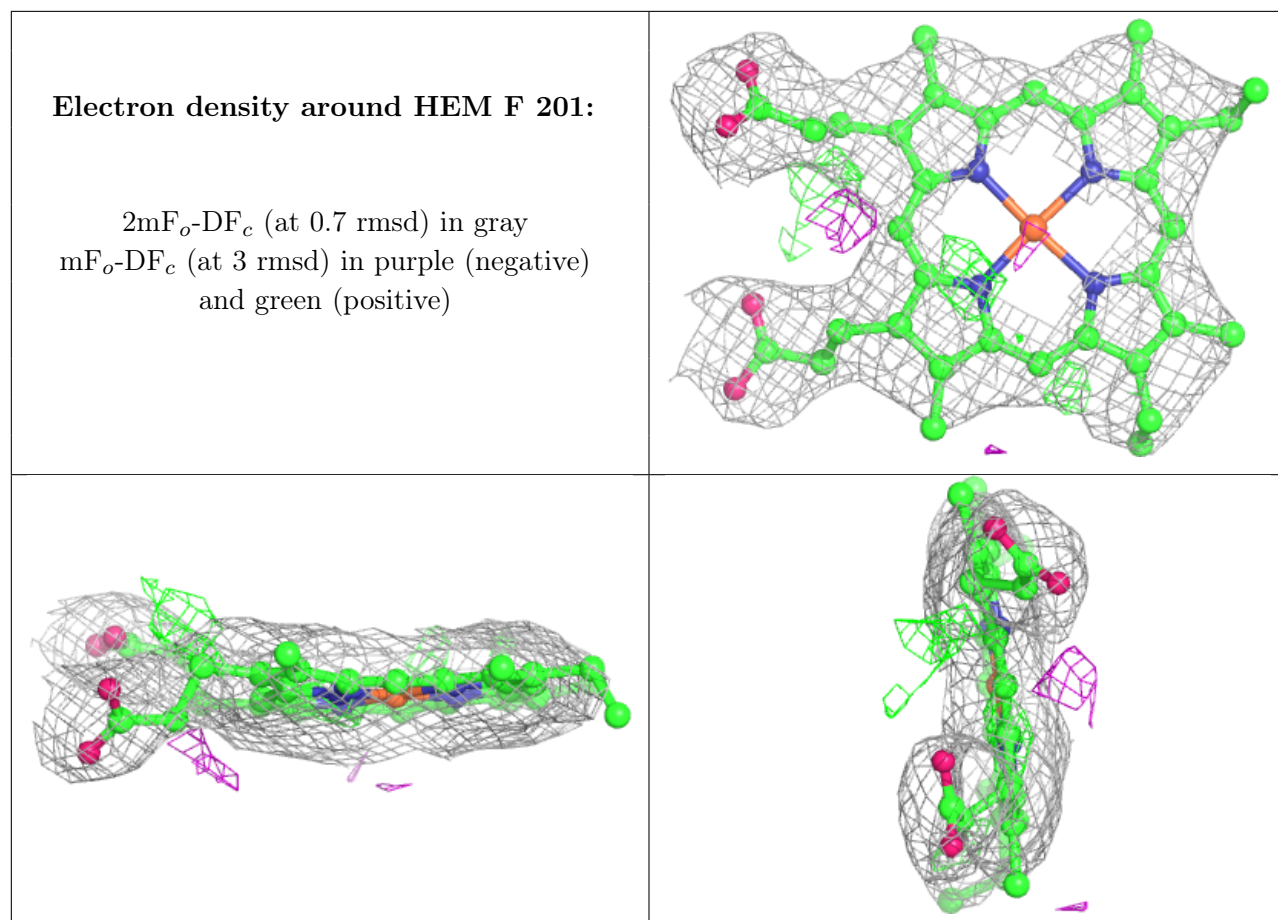
**Electron density around HEM D 201:**

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



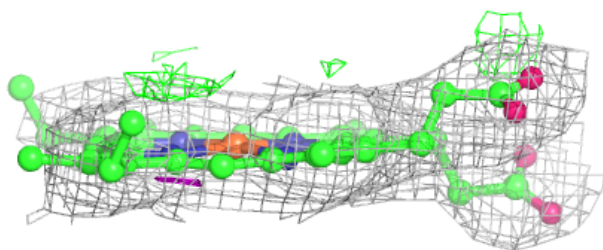
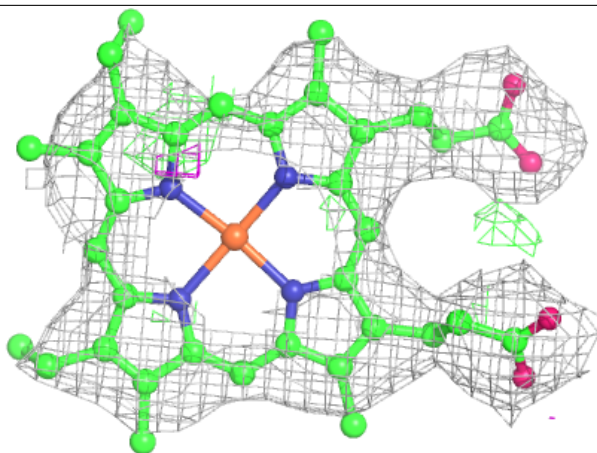


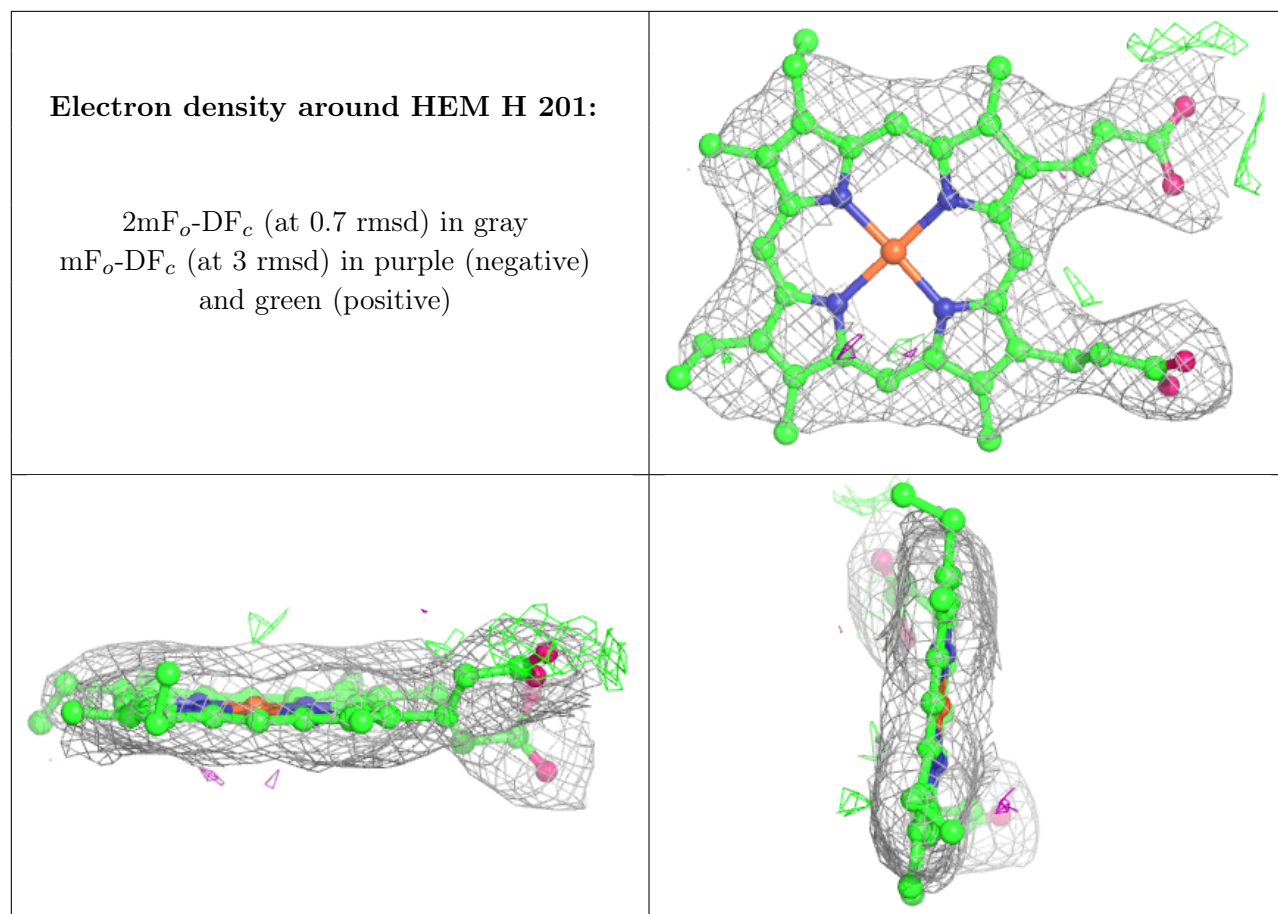




**Electron density around HEM G 201:**

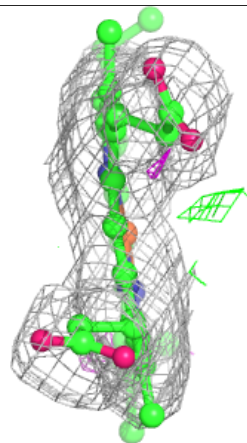
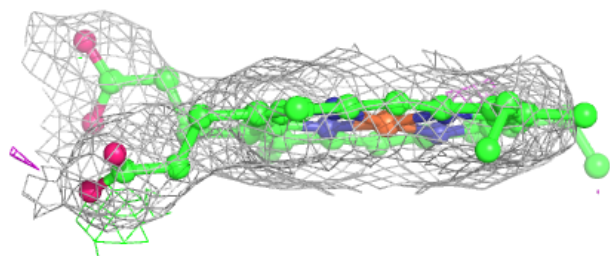
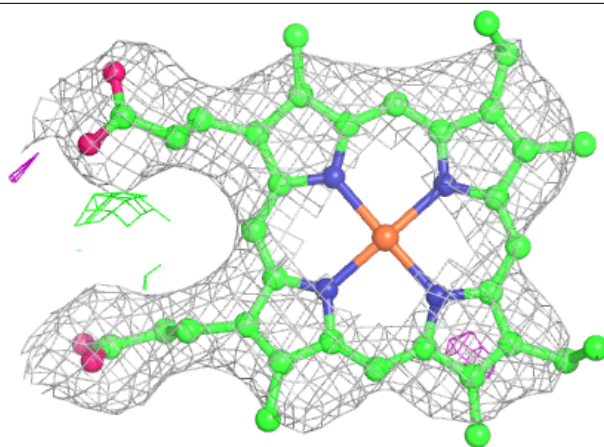
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



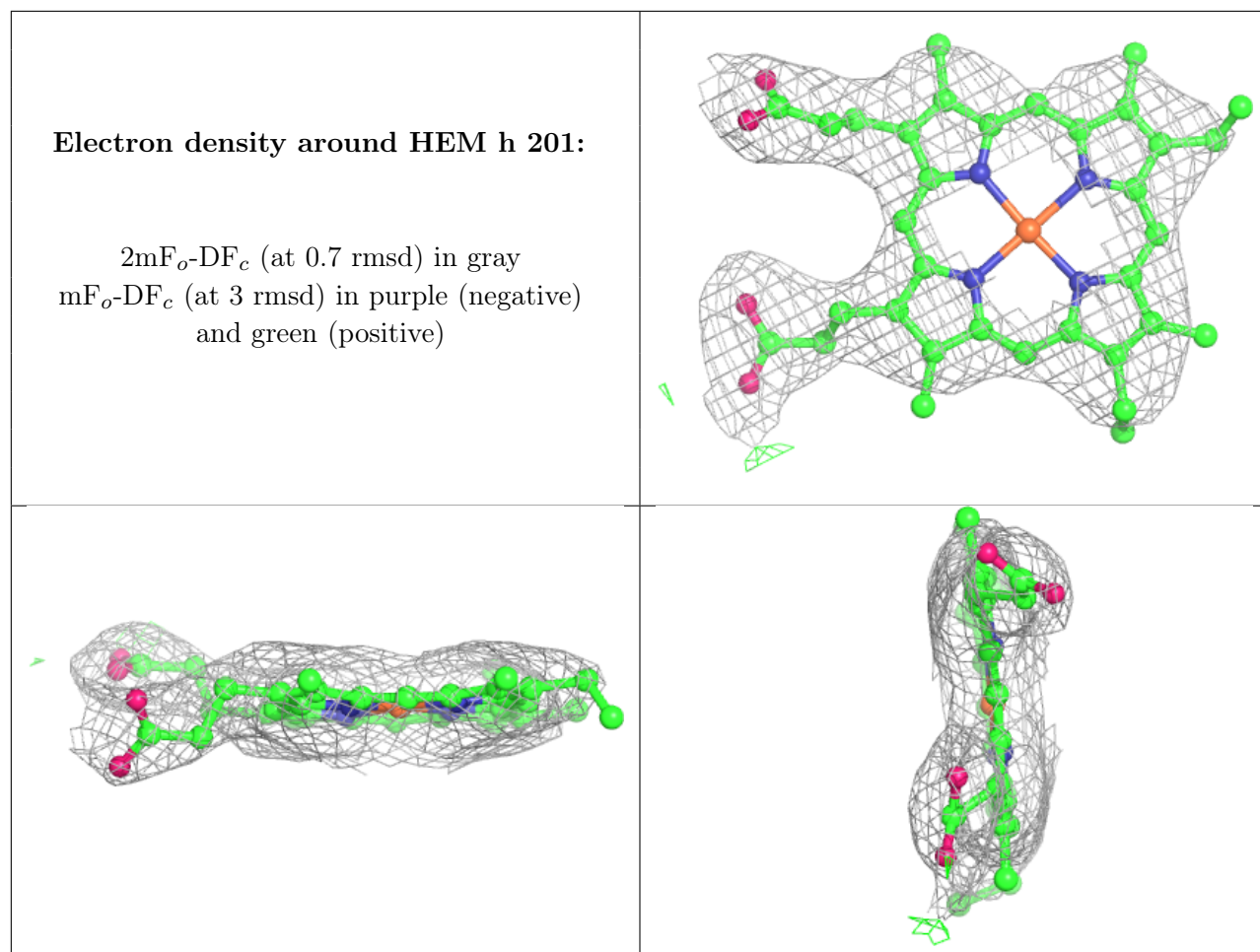


**Electron density around HEM I 201:**

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

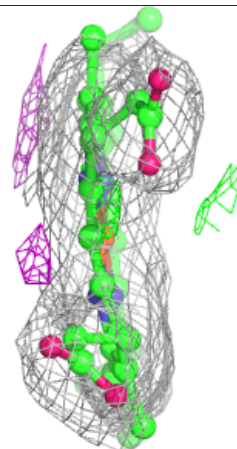
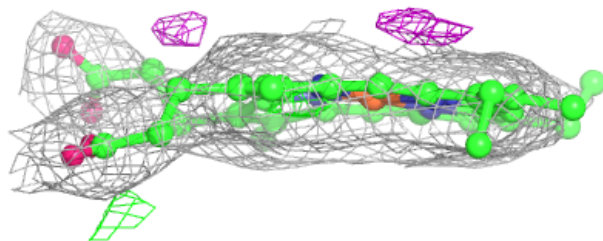
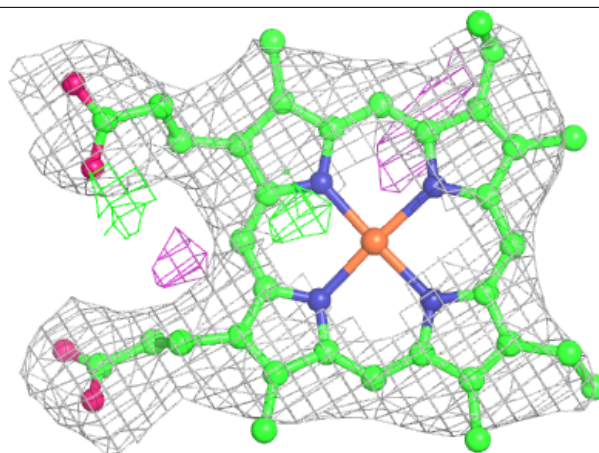


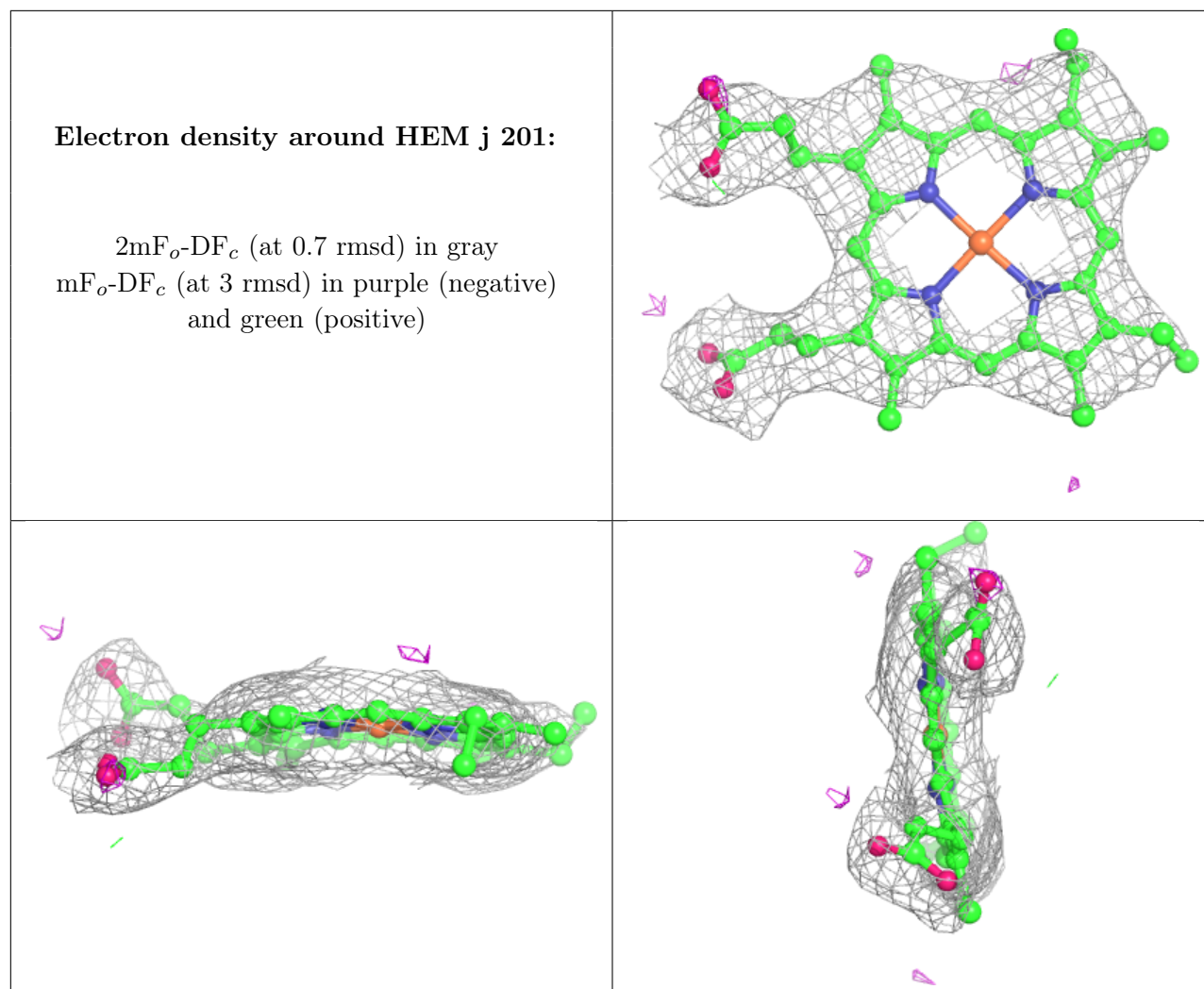




**Electron density around HEM U 201:**

$2mF_o-DF_c$  (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.