



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 12, 2023 – 12:48 AM EDT

PDB ID : 4LMX  
Title : Light harvesting complex PE555 from the cryptophyte *Hemiselmis andersenii*  
CCMP644  
Authors : Harrop, S.J.; Wilk, K.E.; Curmi, P.M.G.  
Deposited on : 2013-07-11  
Resolution : 1.80 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 : 2.35.1  
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) : 2.35.1

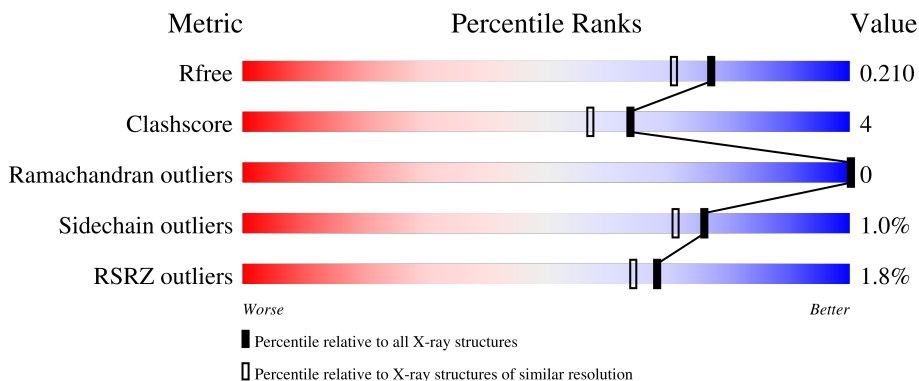
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

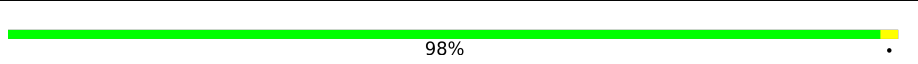
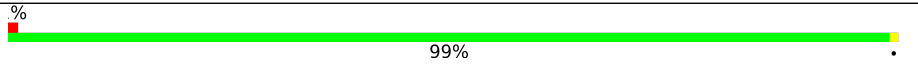
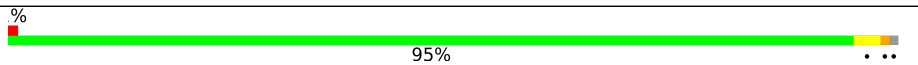
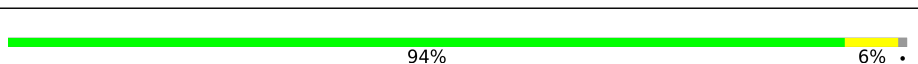
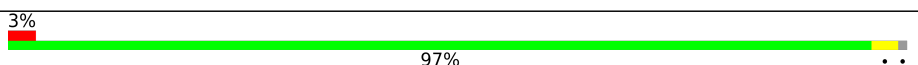
The reported resolution of this entry is 1.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	62	 98%
2	B	177	 99%
2	D	177	 95%
2	F	177	 94% 6%
2	H	177	 97%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	J	177	<p>2% 97% ..</p>
2	L	177	<p>4% 96% ..</p>
3	C	67	<p>% 91% 7% .</p>
4	E	74	<p>3% 89% 9% .</p>
4	G	74	<p>4% 86% 5% 5% .</p>
4	I	74	<p>93% 5% .</p>
4	K	74	<p>3% 86% 12% .</p>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 14008 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 cryptophyte phycoerythrin (alpha-2 chain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	62	466	285	80	94	7	0	1	0

- Molecule 2 is a protein called cryptophyte phycoerythrin (beta chain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	177	1292	800	220	262	10	0	2	0
2	D	175	1267	784	217	257	9	0	0	0
2	F	176	1284	795	220	260	9	0	1	0
2	H	176	1284	795	220	260	9	0	1	0
2	J	175	1294	798	222	265	9	0	4	0
2	L	175	1276	789	219	259	9	0	1	0

- Molecule 3 is a protein called cryptophyte phycoerythrin (alpha-1 chain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	66	480	296	85	94	5	0	0	0

- Molecule 4 is a protein called cryptophyte phycoerythrin (alpha-1/alpha-2 chain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	E	66	536	331	94	105	6	0	7	0
4	G	65	530	328	93	103	6	0	7	0

*Continued on next page...*

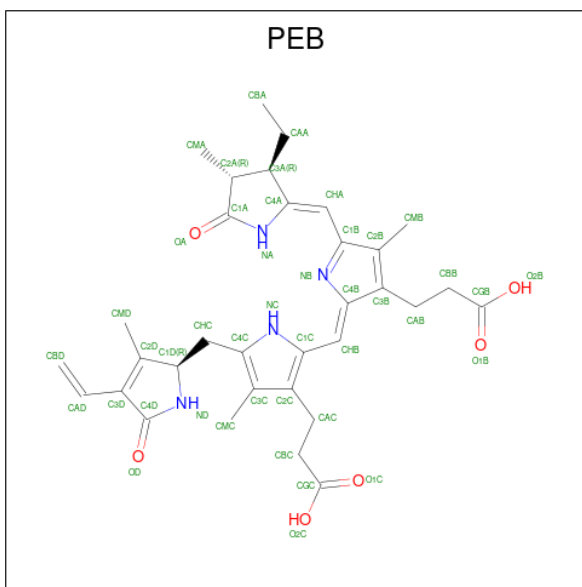
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	I	66	Total	C	N	O	S	0	7	0
			536	331	94	105	6			
4	K	66	Total	C	N	O	S	0	7	0
			536	331	94	105	6			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	44	LYS	GLY	microheterogeneity	PDB 4LMX
E	45	MET	PHE	microheterogeneity	PDB 4LMX
E	51	THR	ALA	microheterogeneity	PDB 4LMX
E	58	ASN	GLY	microheterogeneity	PDB 4LMX
E	59	THR	ILE	microheterogeneity	PDB 4LMX
E	61	LEU	ARG	microheterogeneity	PDB 4LMX
E	62	LEU	PHE	microheterogeneity	PDB 4LMX
G	44	LYS	GLY	microheterogeneity	PDB 4LMX
G	45	MET	PHE	microheterogeneity	PDB 4LMX
G	51	THR	ALA	microheterogeneity	PDB 4LMX
G	58	ASN	GLY	microheterogeneity	PDB 4LMX
G	59	THR	ILE	microheterogeneity	PDB 4LMX
G	61	LEU	ARG	microheterogeneity	PDB 4LMX
G	62	LEU	PHE	microheterogeneity	PDB 4LMX
I	44	LYS	GLY	microheterogeneity	PDB 4LMX
I	45	MET	PHE	microheterogeneity	PDB 4LMX
I	51	THR	ALA	microheterogeneity	PDB 4LMX
I	58	ASN	GLY	microheterogeneity	PDB 4LMX
I	59	THR	ILE	microheterogeneity	PDB 4LMX
I	61	LEU	ARG	microheterogeneity	PDB 4LMX
I	62	LEU	PHE	microheterogeneity	PDB 4LMX
K	44	LYS	GLY	microheterogeneity	PDB 4LMX
K	45	MET	PHE	microheterogeneity	PDB 4LMX
K	51	THR	ALA	microheterogeneity	PDB 4LMX
K	58	ASN	GLY	microheterogeneity	PDB 4LMX
K	59	THR	ILE	microheterogeneity	PDB 4LMX
K	61	LEU	ARG	microheterogeneity	PDB 4LMX
K	62	LEU	PHE	microheterogeneity	PDB 4LMX

- Molecule 5 is PHYCOERYTHROBILIN (three-letter code: PEB) (formula: C<sub>33</sub>H<sub>40</sub>N<sub>4</sub>O<sub>6</sub>).



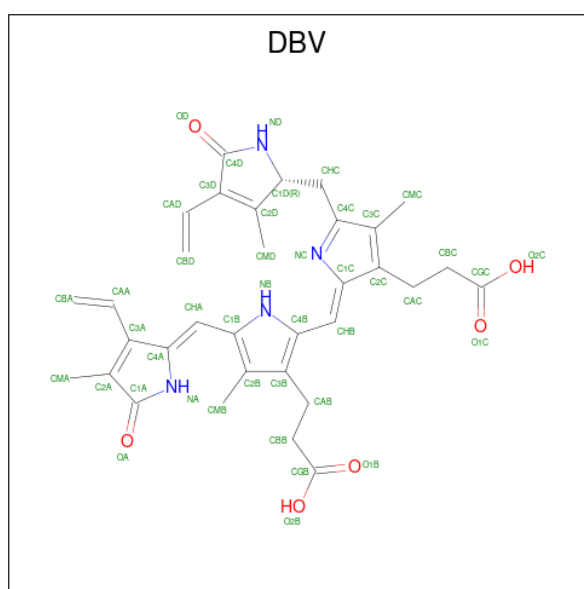
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			43	33	4	6		
5	B	1	Total	C	N	O	0	0
			43	33	4	6		
5	B	1	Total	C	N	O	0	0
			43	33	4	6		
5	C	1	Total	C	N	O	0	0
			43	33	4	6		
5	D	1	Total	C	N	O	0	0
			43	33	4	6		
5	D	1	Total	C	N	O	0	0
			43	33	4	6		
5	E	1	Total	C	N	O	0	0
			43	33	4	6		
5	F	1	Total	C	N	O	0	0
			43	33	4	6		
5	F	1	Total	C	N	O	0	0
			43	33	4	6		
5	G	1	Total	C	N	O	0	0
			43	33	4	6		
5	H	1	Total	C	N	O	0	0
			43	33	4	6		
5	H	1	Total	C	N	O	0	0
			43	33	4	6		
5	I	1	Total	C	N	O	0	0
			43	33	4	6		
5	J	1	Total	C	N	O	0	0
			43	33	4	6		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	J	1	Total	C	N	O	0	0
			43	33	4	6		
5	K	1	Total	C	N	O	0	0
			43	33	4	6		
5	L	1	Total	C	N	O	0	0
			43	33	4	6		
5	L	1	Total	C	N	O	0	0
			43	33	4	6		

- Molecule 6 is 15,16-DIHYDROBILIVERDIN (three-letter code: DBV) (formula: C<sub>33</sub>H<sub>36</sub>N<sub>4</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total	C	N	O	0	0
			43	33	4	6		
6	D	1	Total	C	N	O	0	0
			43	33	4	6		
6	F	1	Total	C	N	O	0	0
			43	33	4	6		
6	H	1	Total	C	N	O	0	0
			43	33	4	6		
6	J	1	Total	C	N	O	0	0
			43	33	4	6		
6	L	1	Total	C	N	O	0	0
			43	33	4	6		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	111	Total O 111 111	0	0
7	B	282	Total O 282 282	0	0
7	C	136	Total O 136 136	0	0
7	D	248	Total O 248 248	0	0
7	E	129	Total O 129 129	0	0
7	F	274	Total O 274 274	0	0
7	G	97	Total O 97 97	0	0
7	H	254	Total O 254 254	0	0
7	I	116	Total O 116 116	0	0
7	J	271	Total O 271 271	0	0
7	K	104	Total O 104 104	0	0
7	L	173	Total O 173 173	0	0



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

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

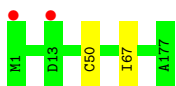
- Molecule 1: cryptophyte phycoerythrin (alpha-2 chain)

Chain A:  98%



- Molecule 2: cryptophyte phycoerythrin (beta chain)

Chain B:  99%



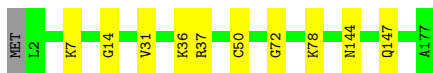
- Molecule 2: cryptophyte phycoerythrin (beta chain)

Chain D:  95%



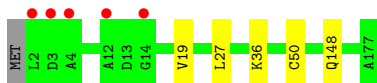
- Molecule 2: cryptophyte phycoerythrin (beta chain)

Chain F:  94%

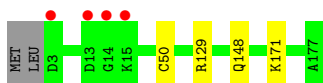


- Molecule 2: cryptophyte phycoerythrin (beta chain)

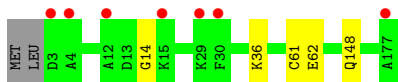
Chain H:  97%



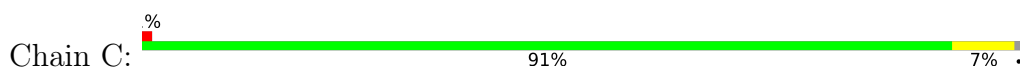
- Molecule 2: cryptophyte phycoerythrin (beta chain)



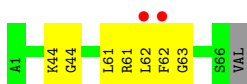
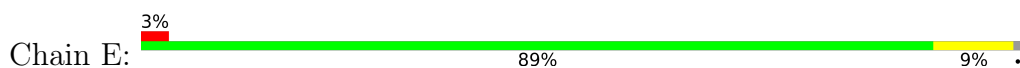
- Molecule 2: cryptophyte phycoerythrin (beta chain)



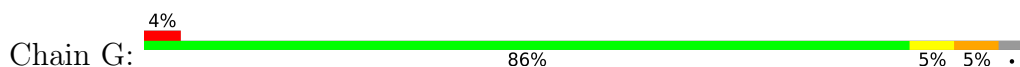
- Molecule 3: cryptophyte phycoerythrin (alpha-1 chain)



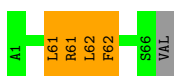
- Molecule 4: cryptophyte phycoerythrin (alpha-1/alpha-2 chain)



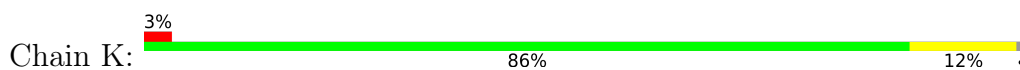
- Molecule 4: cryptophyte phycoerythrin (alpha-1/alpha-2 chain)



- Molecule 4: cryptophyte phycoerythrin (alpha-1/alpha-2 chain)



- Molecule 4: cryptophyte phycoerythrin (alpha-1/alpha-2 chain)



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.07Å 76.74Å 142.59Å 90.00° 92.86° 90.00°	Depositor
Resolution (Å)	18.64 – 1.80 18.64 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (18.64-1.80) 96.9 (18.64-1.80)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.26 (at 1.80Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.2_867)	Depositor
R, $R_{free}$	0.158 , 0.214 0.155 , 0.210	Depositor DCC
$R_{free}$ test set	6624 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.4	Xtrriage
Anisotropy	0.537	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 53.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.016 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14008	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.03 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.1971e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: DBV, LYZ, PEB

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	0.35	0/458	0.53	0/607
2	B	0.34	0/1308	0.47	0/1764
2	D	0.34	0/1280	0.47	0/1727
2	F	0.34	0/1297	0.46	0/1750
2	H	0.34	0/1297	0.45	0/1750
2	J	0.34	0/1307	0.49	0/1763
2	L	0.30	0/1289	0.44	0/1739
3	C	0.34	0/474	0.54	0/628
4	E	0.35	0/522	0.56	0/678
4	G	0.56	2/516 (0.4%)	0.55	0/670
4	I	0.34	0/522	0.63	1/678 (0.1%)
4	K	0.38	0/522	0.52	0/678
All	All	0.35	2/10792 (0.0%)	0.49	1/14432 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	58[B]	GLY	C-N	-7.35	1.17	1.34
4	G	44[B]	GLY	C-N	-7.28	1.17	1.34

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	I	61[B]	ARG	NE-CZ-NH1	5.84	123.22	120.30

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	466	0	473	1	0
2	B	1292	0	1305	1	0
2	D	1267	0	1273	5	0
2	F	1284	0	1288	6	0
2	H	1284	0	1287	8	0
2	J	1294	0	1289	3	0
2	L	1276	0	1277	10	0
3	C	480	0	480	4	0
4	E	536	0	531	8	0
4	G	530	0	526	7	0
4	I	536	0	530	6	0
4	K	536	0	532	9	0
5	A	43	0	37	1	0
5	B	86	0	74	2	0
5	C	43	0	37	0	0
5	D	86	0	74	2	0
5	E	43	0	37	2	0
5	F	86	0	74	4	0
5	G	43	0	37	2	0
5	H	86	0	74	2	0
5	I	43	0	37	2	0
5	J	86	0	74	2	0
5	K	43	0	37	0	0
5	L	86	0	74	3	0
6	B	43	0	32	1	0
6	D	43	0	32	5	0
6	F	43	0	32	1	0
6	H	43	0	32	2	0
6	J	43	0	32	0	0
6	L	43	0	32	0	0
7	A	111	0	0	0	0
7	B	282	0	0	0	0
7	C	136	0	0	1	0
7	D	248	0	0	0	0
7	E	129	0	0	1	0
7	F	274	0	0	3	1
7	G	97	0	0	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	H	254	0	0	0	0
7	I	116	0	0	0	1
7	J	271	0	0	3	0
7	K	104	0	0	4	0
7	L	173	0	0	0	0
All	All	14008	0	11649	70	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

The worst 5 of 70 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)
4:K:63:GLY:N	2:L:148[A]:GLN:NE2	1.67	1.39
4:G:63:GLY:N	2:H:148[A]:GLN:NE2	1.73	1.34
4:I:62[A]:LEU:HD12	2:L:148[A]:GLN:HG3	1.60	0.82
4:E:62[A]:LEU:HD12	2:H:148[A]:GLN:NE2	2.03	0.73
2:J:129:ARG:NH1	7:J:458:HOH:O	2.21	0.73

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:F:504:HOH:O	7:I:279:HOH:O[2_756]	2.11	0.09

## 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	60/62 (97%)	60 (100%)	0	0	100   100
2	B	177/177 (100%)	175 (99%)	2 (1%)	0	100   100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	173/177 (98%)	171 (99%)	2 (1%)	0	100	100
2	F	175/177 (99%)	172 (98%)	3 (2%)	0	100	100
2	H	175/177 (99%)	172 (98%)	3 (2%)	0	100	100
2	J	177/177 (100%)	175 (99%)	2 (1%)	0	100	100
2	L	174/177 (98%)	172 (99%)	2 (1%)	0	100	100
3	C	63/67 (94%)	63 (100%)	0	0	100	100
4	E	56/74 (76%)	56 (100%)	0	0	100	100
4	G	55/74 (74%)	55 (100%)	0	0	100	100
4	I	56/74 (76%)	56 (100%)	0	0	100	100
4	K	56/74 (76%)	55 (98%)	1 (2%)	0	100	100
All	All	1397/1487 (94%)	1382 (99%)	15 (1%)	0	100	100

There are no Ramachandran outliers to report.

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

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	49/48 (102%)	49 (100%)	0	100	100
2	B	142/140 (101%)	141 (99%)	1 (1%)	84	81
2	D	138/140 (99%)	136 (99%)	2 (1%)	67	59
2	F	140/140 (100%)	138 (99%)	2 (1%)	67	59
2	H	140/140 (100%)	139 (99%)	1 (1%)	84	81
2	J	142/140 (101%)	141 (99%)	1 (1%)	84	81
2	L	139/140 (99%)	137 (99%)	2 (1%)	67	59
3	C	47/48 (98%)	47 (100%)	0	100	100
4	E	54/55 (98%)	54 (100%)	0	100	100
4	G	53/55 (96%)	53 (100%)	0	100	100
4	I	54/55 (98%)	52 (96%)	2 (4%)	34	19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	K	54/55 (98%)	54 (100%)	0	100	100
All	All	1152/1156 (100%)	1141 (99%)	11 (1%)	76	71

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	I	62[B]	PHE
2	J	50	CYS
2	L	62	GLU
2	L	61	CYS
2	F	144	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

6 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
1	LYZ	A	4	1	7,9,10	0.48	0	4,10,12	0.62	0
4	LYZ	I	4	4	7,9,10	0.48	0	4,10,12	0.77	0
4	LYZ	E	4	4	7,9,10	0.52	0	4,10,12	0.67	0
3	LYZ	C	4	3	7,9,10	0.48	0	4,10,12	0.60	0
4	LYZ	G	4	4	7,9,10	0.52	0	4,10,12	0.61	0
4	LYZ	K	4	4	7,9,10	0.51	0	4,10,12	0.59	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral



centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	LYZ	A	4	1	-	2/8/9/11	-
4	LYZ	I	4	4	-	0/8/9/11	-
4	LYZ	E	4	4	-	0/8/9/11	-
3	LYZ	C	4	3	-	0/8/9/11	-
4	LYZ	G	4	4	-	0/8/9/11	-
4	LYZ	K	4	4	-	0/8/9/11	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	4	LYZ	CG-CD-CE-NZ
1	A	4	LYZ	OH-CD-CE-NZ

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	K	4	LYZ	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

24 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	PEB	E	101	4	43,46,46	2.26	7 (16%)	45,67,67	1.44	7 (15%)
5	PEB	L	203	2	43,46,46	2.29	6 (13%)	45,67,67	1.73	9 (20%)
5	PEB	J	203	2	43,46,46	2.30	5 (11%)	45,67,67	1.67	11 (24%)
5	PEB	H	202	2	43,46,46	2.28	5 (11%)	45,67,67	1.47	7 (15%)
5	PEB	I	101	4	43,46,46	2.30	4 (9%)	45,67,67	1.48	7 (15%)
5	PEB	K	101	4	43,46,46	2.34	6 (13%)	45,67,67	1.54	8 (17%)
6	DBV	H	201	2	42,46,46	2.79	10 (23%)	42,67,67	1.57	6 (14%)
5	PEB	B	202	2	43,46,46	2.31	6 (13%)	45,67,67	1.44	8 (17%)
6	DBV	B	201	2	42,46,46	2.82	11 (26%)	42,67,67	1.59	8 (19%)
5	PEB	A	101	1	43,46,46	2.16	6 (13%)	45,67,67	1.51	8 (17%)
5	PEB	H	203	2	43,46,46	2.37	6 (13%)	45,67,67	1.76	11 (24%)
5	PEB	C	101	3	43,46,46	2.21	6 (13%)	45,67,67	1.37	6 (13%)
5	PEB	F	202	2	43,46,46	2.26	6 (13%)	45,67,67	1.36	5 (11%)
6	DBV	D	201	2	42,46,46	2.74	10 (23%)	42,67,67	1.54	8 (19%)
6	DBV	F	201	2	42,46,46	2.82	10 (23%)	42,67,67	1.52	6 (14%)
6	DBV	L	201	2	42,46,46	2.87	10 (23%)	42,67,67	1.50	7 (16%)
5	PEB	D	203	2	43,46,46	2.22	6 (13%)	45,67,67	1.69	10 (22%)
6	DBV	J	201	2	42,46,46	2.84	10 (23%)	42,67,67	1.55	5 (11%)
5	PEB	G	101	4	43,46,46	2.27	5 (11%)	45,67,67	1.48	6 (13%)
5	PEB	L	202	2	43,46,46	2.25	6 (13%)	45,67,67	1.50	7 (15%)
5	PEB	J	202	2	43,46,46	2.26	5 (11%)	45,67,67	1.40	5 (11%)
5	PEB	D	202	2	43,46,46	2.27	6 (13%)	45,67,67	1.48	8 (17%)
5	PEB	B	203	2	43,46,46	2.22	5 (11%)	45,67,67	1.62	10 (22%)
5	PEB	F	203	2	43,46,46	2.21	7 (16%)	45,67,67	1.74	8 (17%)

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
5	PEB	E	101	4	-	4/24/74/74	0/4/4/4
5	PEB	L	203	2	-	8/24/74/74	0/4/4/4
5	PEB	J	203	2	-	5/24/74/74	0/4/4/4
5	PEB	H	202	2	-	8/24/74/74	0/4/4/4
5	PEB	I	101	4	-	6/24/74/74	0/4/4/4

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PEB	K	101	4	-	5/24/74/74	0/4/4/4
6	DBV	H	201	2	-	10/26/74/74	0/4/4/4
5	PEB	B	202	2	-	9/24/74/74	0/4/4/4
6	DBV	B	201	2	-	11/26/74/74	0/4/4/4
5	PEB	A	101	1	-	5/24/74/74	0/4/4/4
5	PEB	H	203	2	-	4/24/74/74	0/4/4/4
5	PEB	C	101	3	-	2/24/74/74	0/4/4/4
5	PEB	F	202	2	-	8/24/74/74	0/4/4/4
6	DBV	D	201	2	-	11/26/74/74	0/4/4/4
6	DBV	F	201	2	-	8/26/74/74	0/4/4/4
6	DBV	L	201	2	-	9/26/74/74	0/4/4/4
5	PEB	D	203	2	-	6/24/74/74	0/4/4/4
6	DBV	J	201	2	-	11/26/74/74	0/4/4/4
5	PEB	G	101	4	-	2/24/74/74	0/4/4/4
5	PEB	L	202	2	-	8/24/74/74	0/4/4/4
5	PEB	J	202	2	-	9/24/74/74	0/4/4/4
5	PEB	D	202	2	-	9/24/74/74	0/4/4/4
5	PEB	B	203	2	-	4/24/74/74	0/4/4/4
5	PEB	F	203	2	-	4/24/74/74	0/4/4/4

The worst 5 of 164 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	H	203	PEB	CHB-C4B	12.01	1.45	1.35
6	B	201	DBV	CHB-C1C	11.81	1.45	1.35
6	F	201	DBV	CHB-C1C	11.80	1.45	1.35
6	L	201	DBV	CHB-C1C	11.76	1.44	1.35
5	K	101	PEB	CHB-C4B	11.70	1.44	1.35

The worst 5 of 181 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	L	203	PEB	CHC-C1D-ND	-5.90	107.09	113.95
5	G	101	PEB	CHC-C1D-ND	-4.57	108.64	113.95
6	H	201	DBV	CBA-CAA-C3A	-4.50	105.24	127.62
6	J	201	DBV	CBA-CAA-C3A	-4.34	106.00	127.62
5	F	203	PEB	CHC-C1D-ND	-4.32	108.93	113.95

There are no chirality outliers.

5 of 166 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	101	PEB	NB-C1B-CHA-C4A
5	A	101	PEB	C2B-C1B-CHA-C4A
5	B	202	PEB	C2A-C3A-CAA-CBA
5	B	202	PEB	C4A-C3A-CAA-CBA
5	B	202	PEB	NB-C1B-CHA-C4A

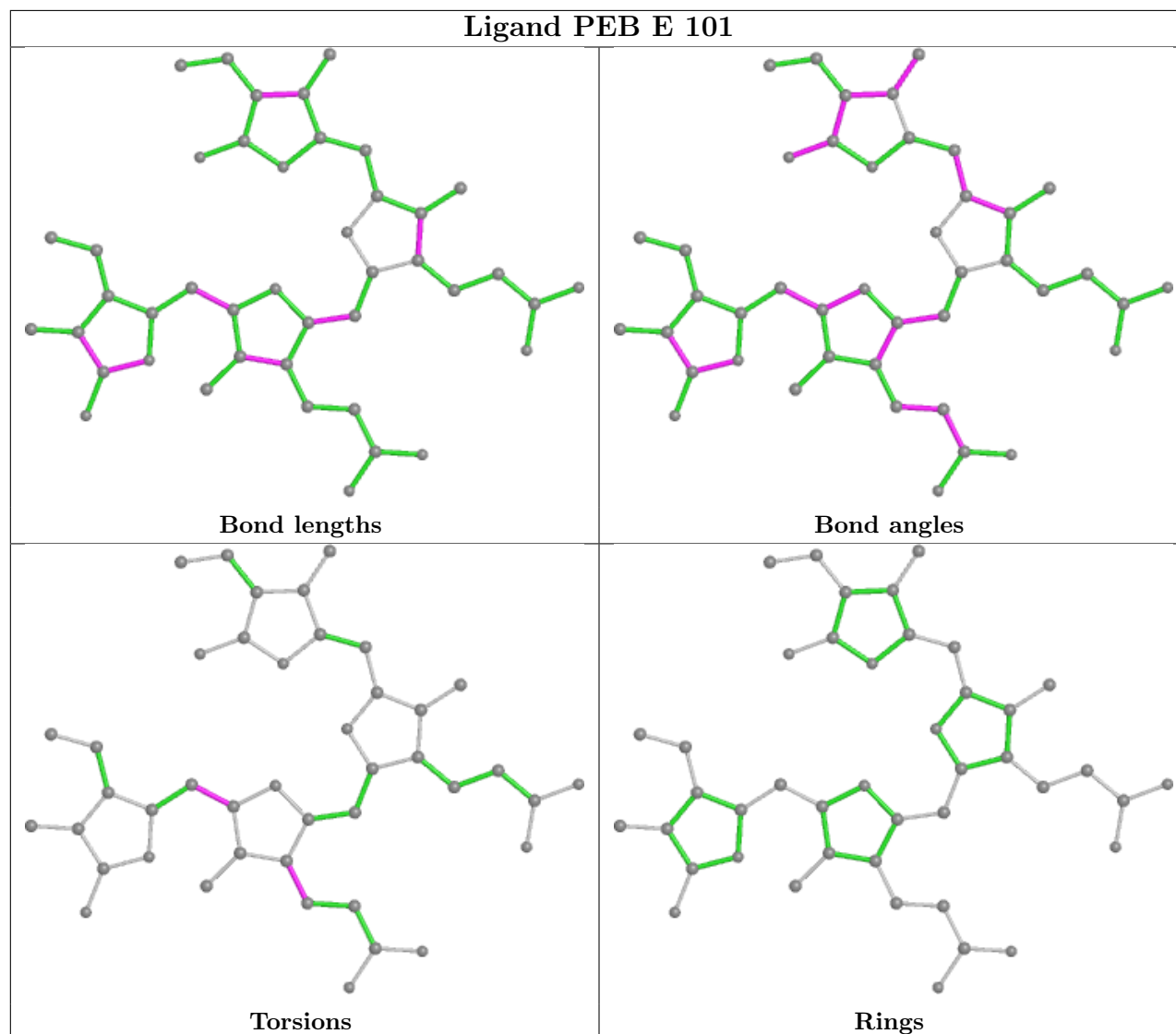
There are no ring outliers.

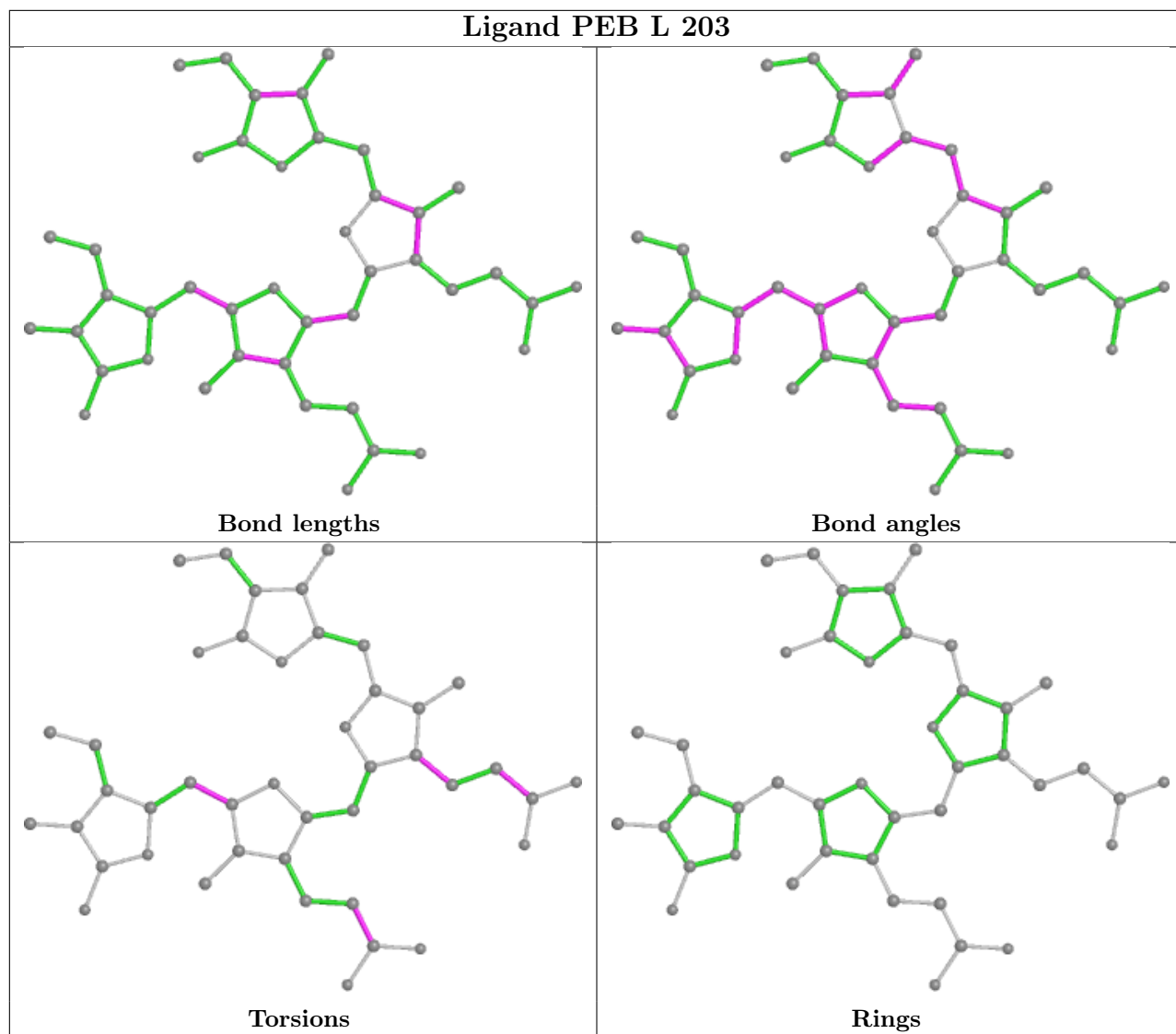
20 monomers are involved in 31 short contacts:

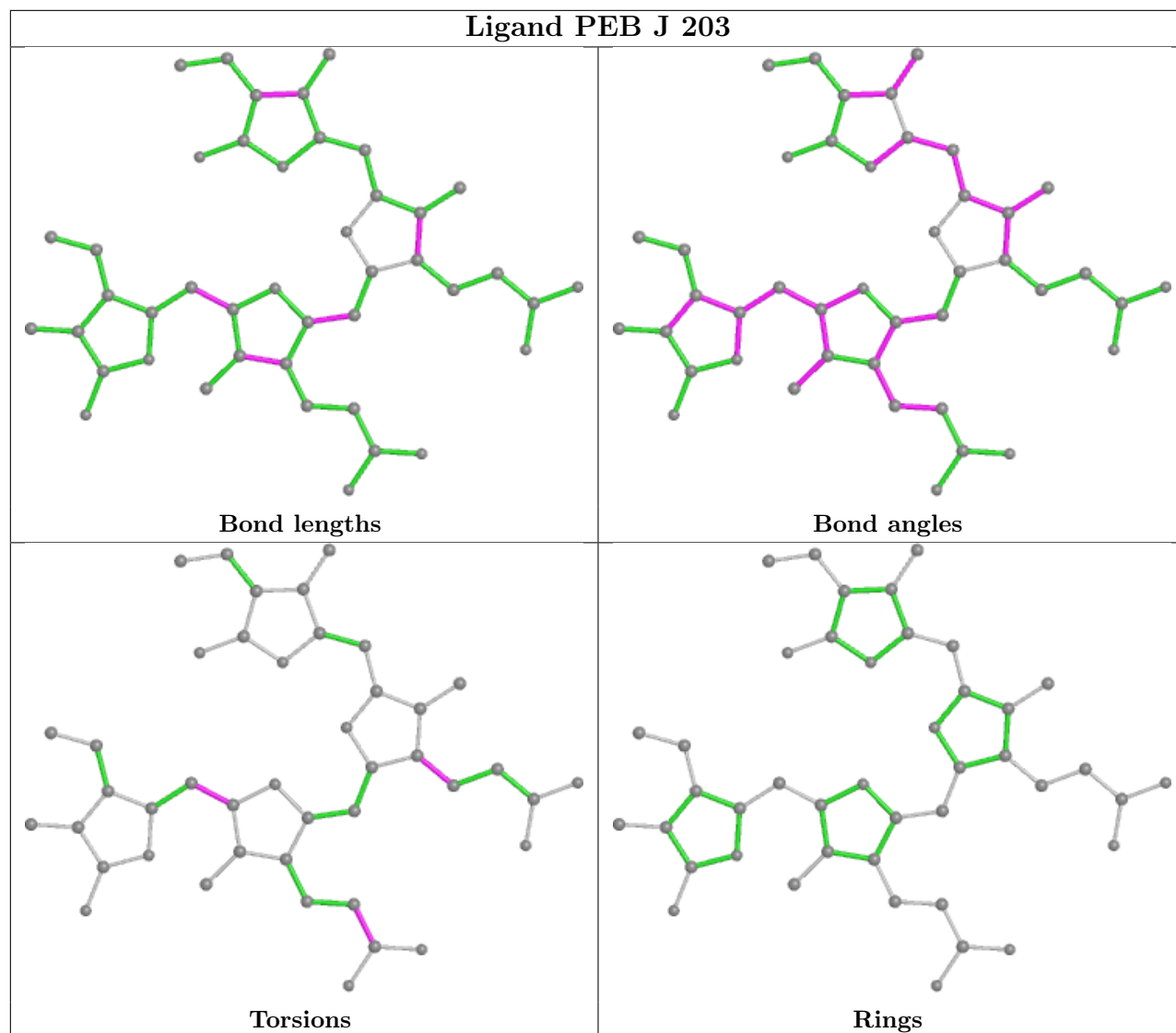
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	101	PEB	2	0
5	L	203	PEB	1	0
5	J	203	PEB	1	0
5	H	202	PEB	1	0
5	I	101	PEB	2	0
6	H	201	DBV	2	0
5	B	202	PEB	1	0
6	B	201	DBV	1	0
5	A	101	PEB	1	0
5	H	203	PEB	1	0
5	F	202	PEB	3	0
6	D	201	DBV	5	0
6	F	201	DBV	1	0
5	D	203	PEB	1	0
5	G	101	PEB	2	0
5	L	202	PEB	2	0
5	J	202	PEB	1	0
5	D	202	PEB	1	0
5	B	203	PEB	1	0
5	F	203	PEB	1	0

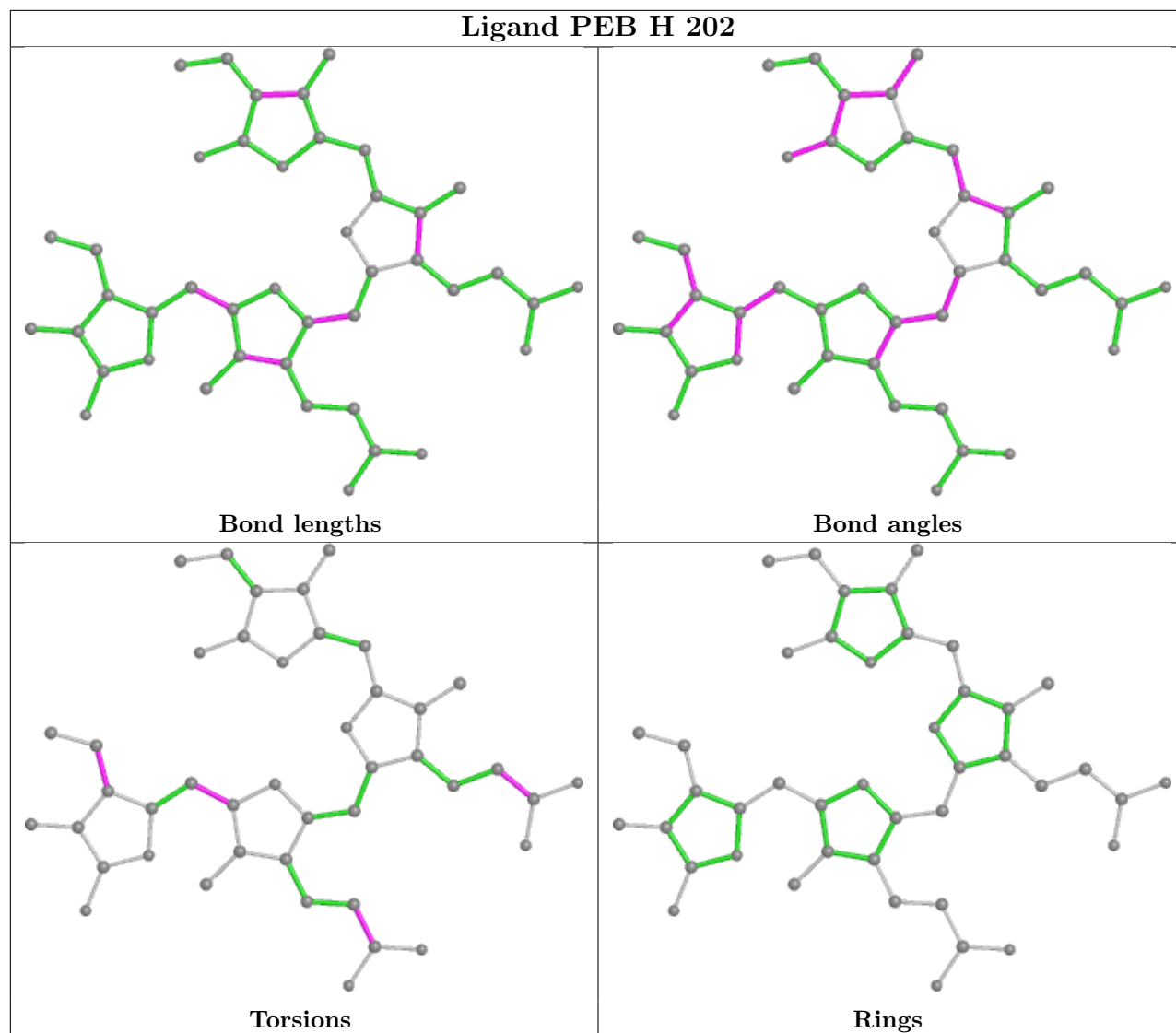
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and

any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

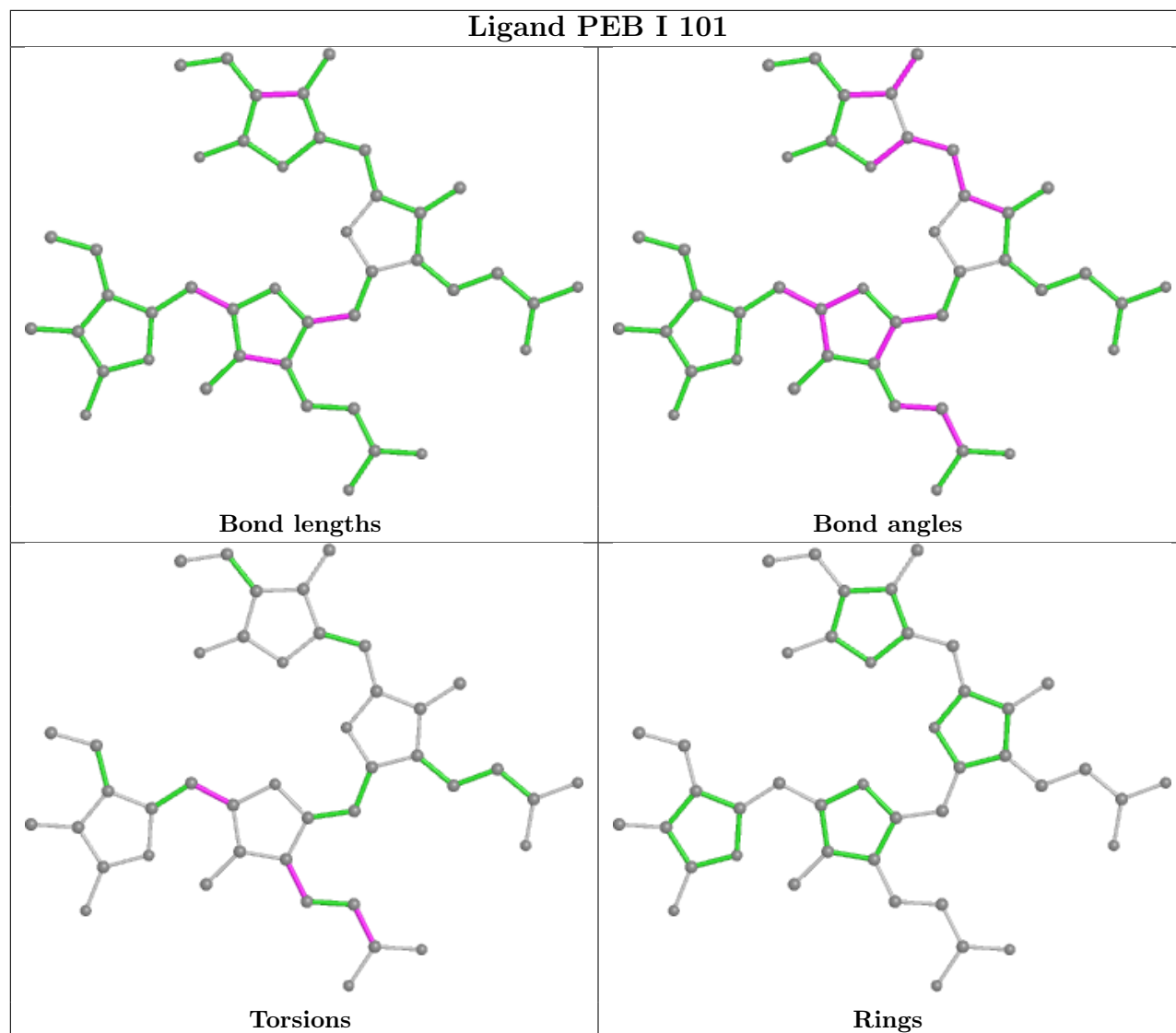


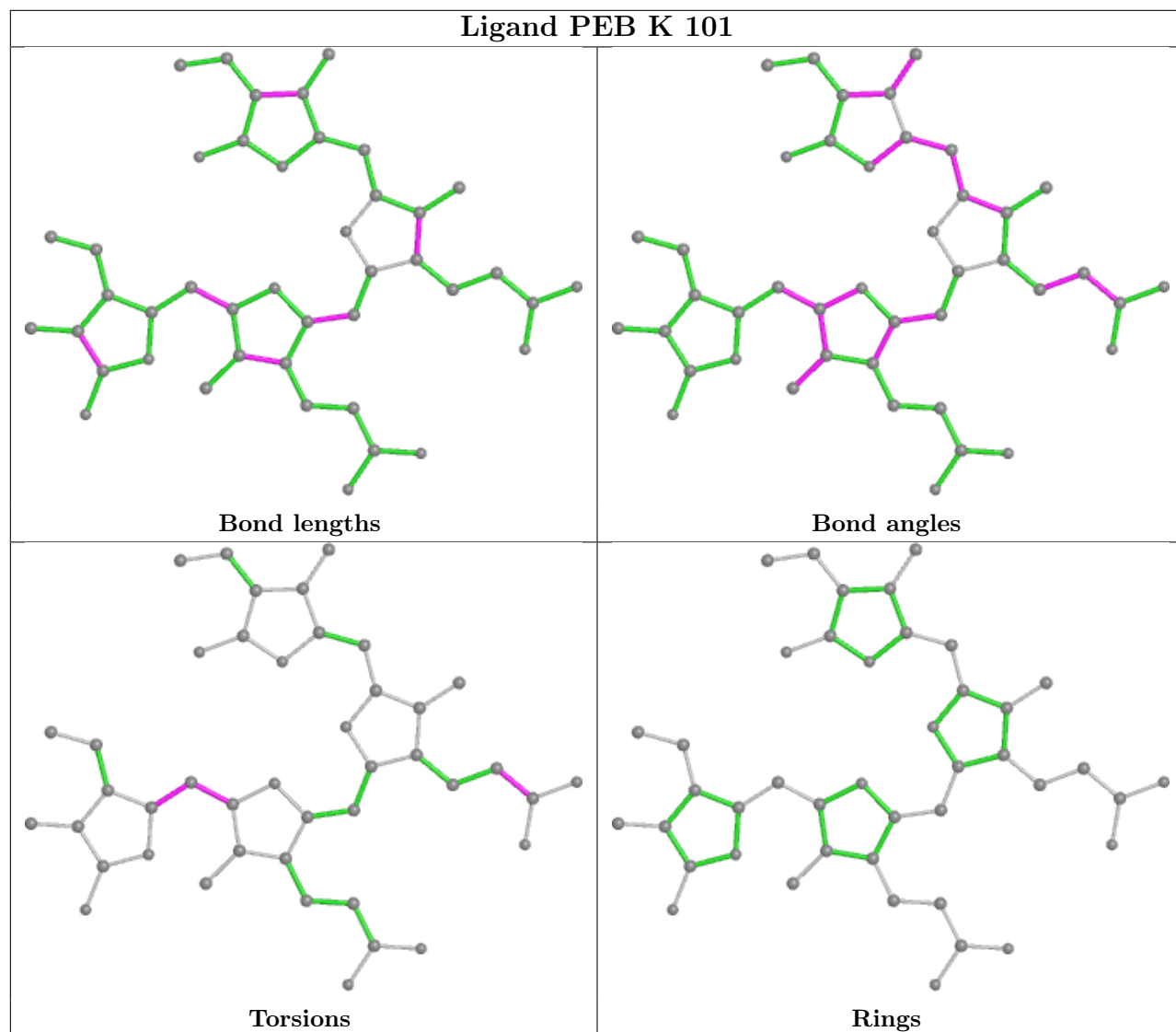


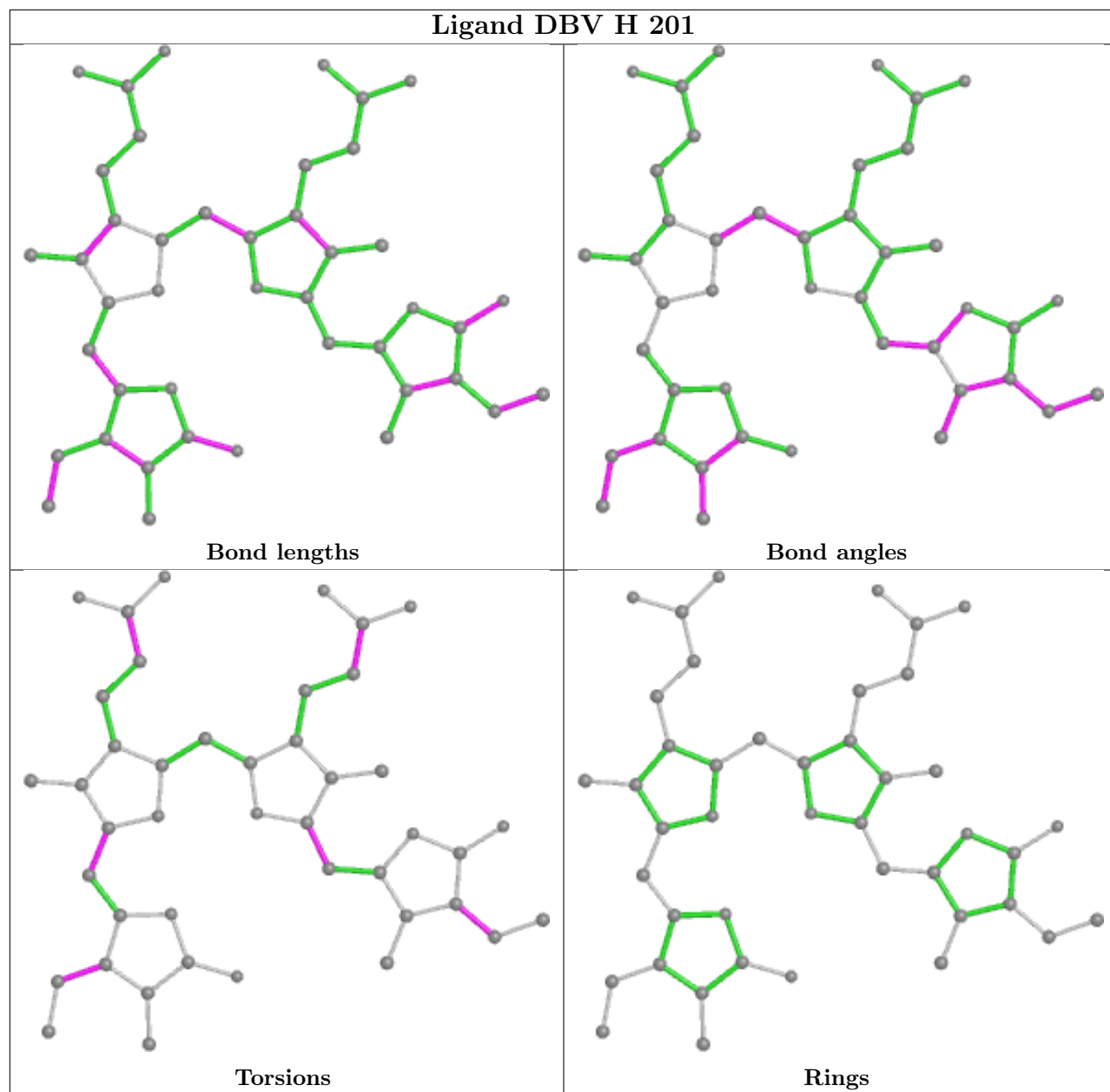


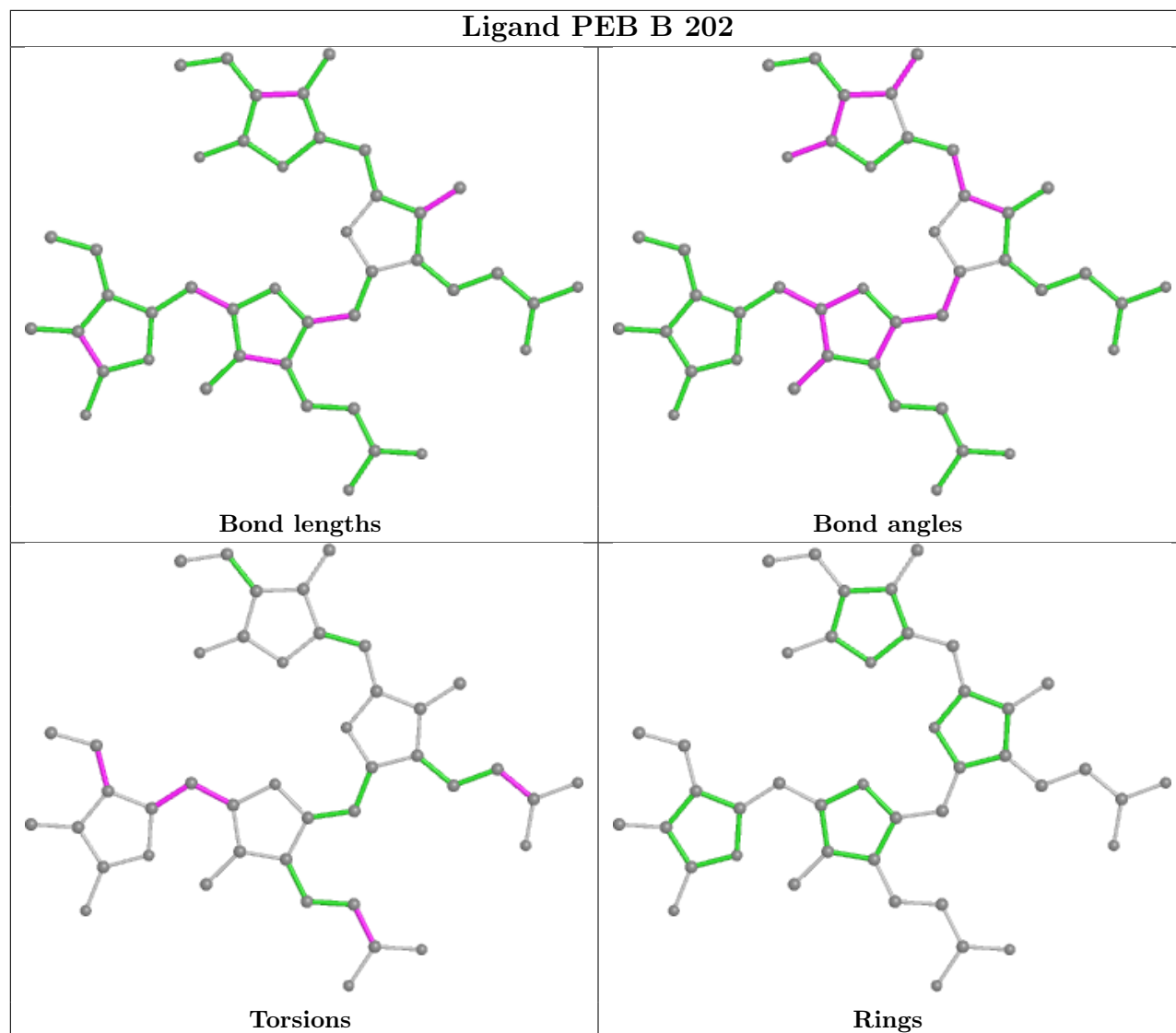


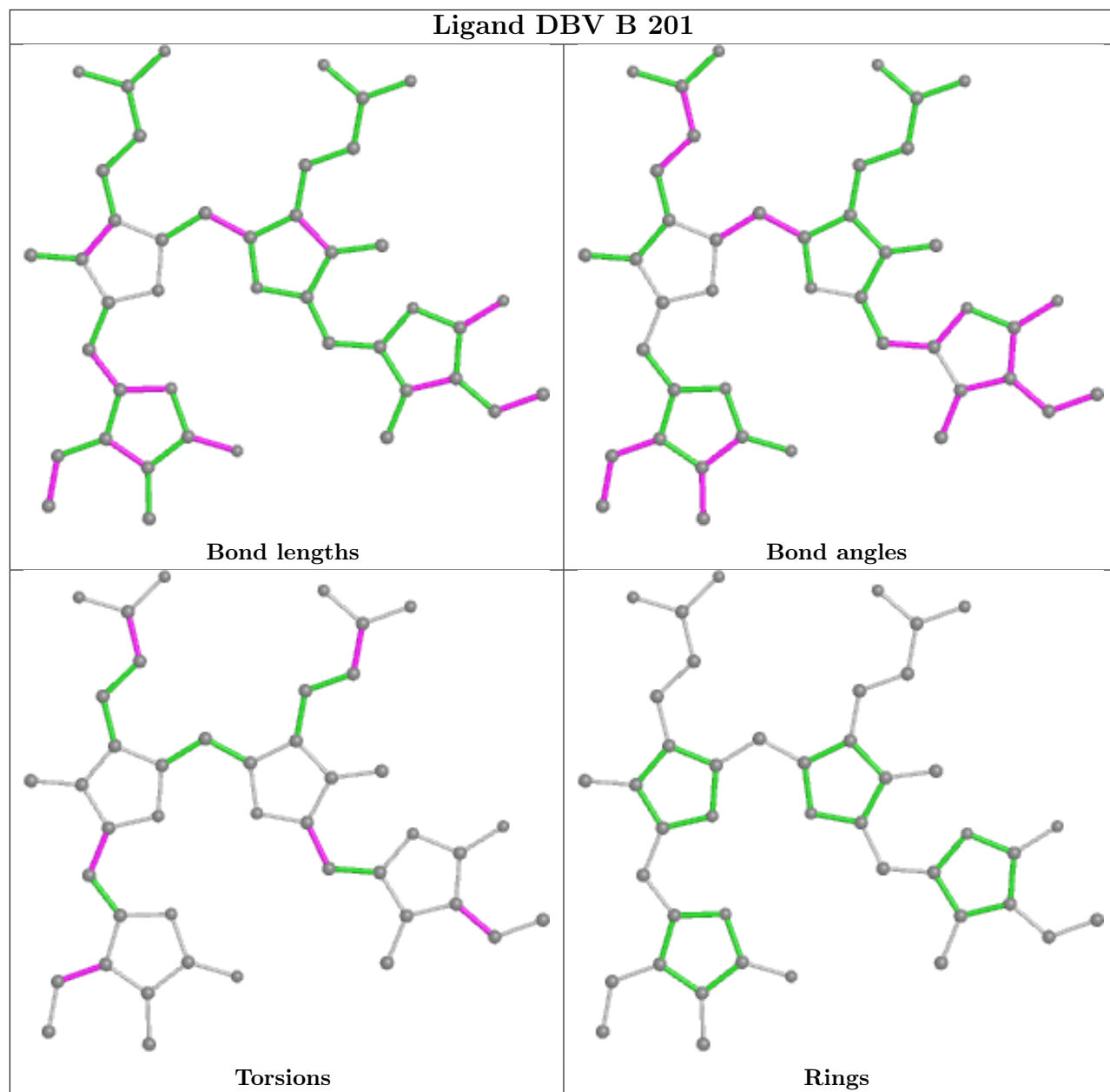


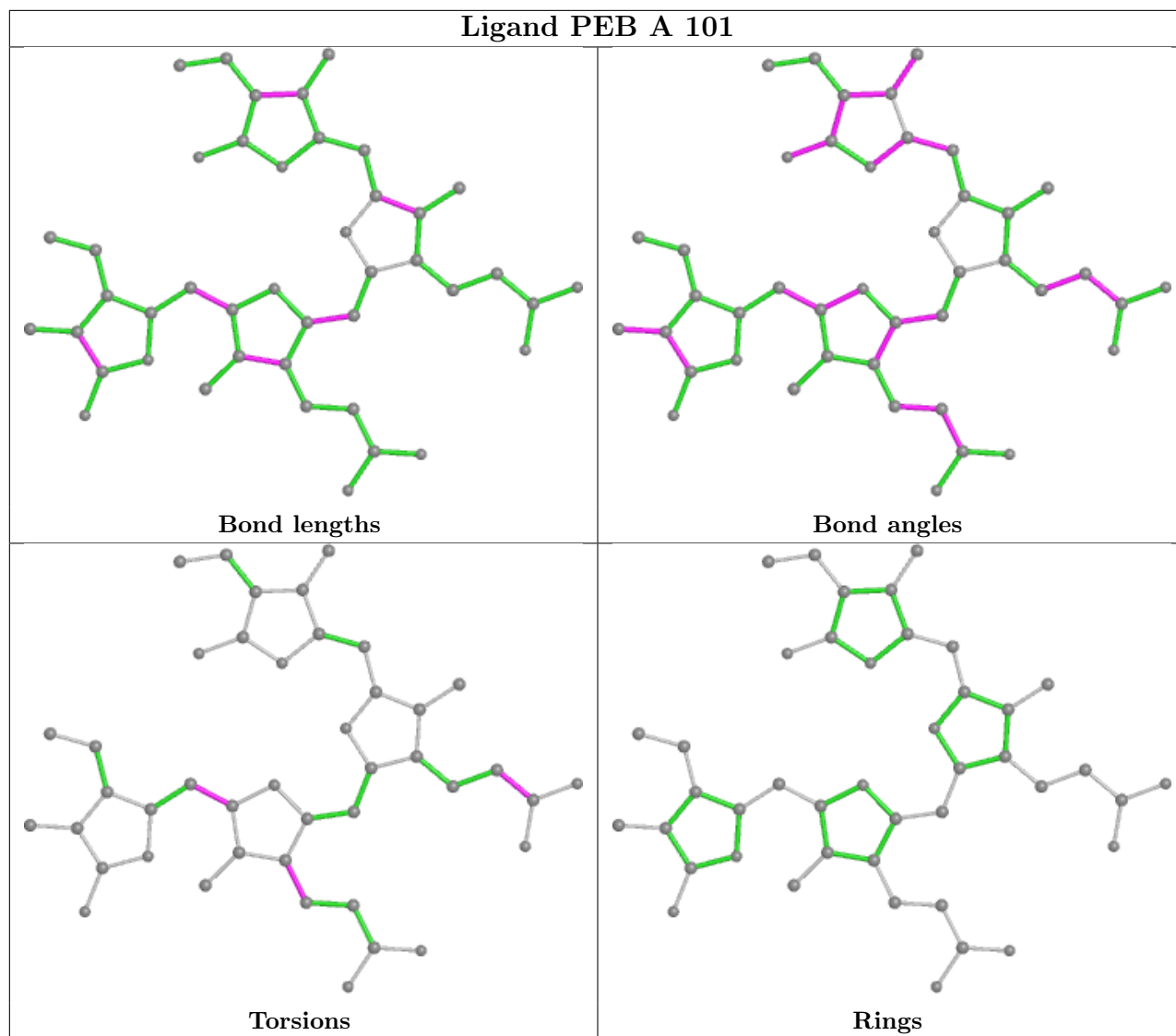


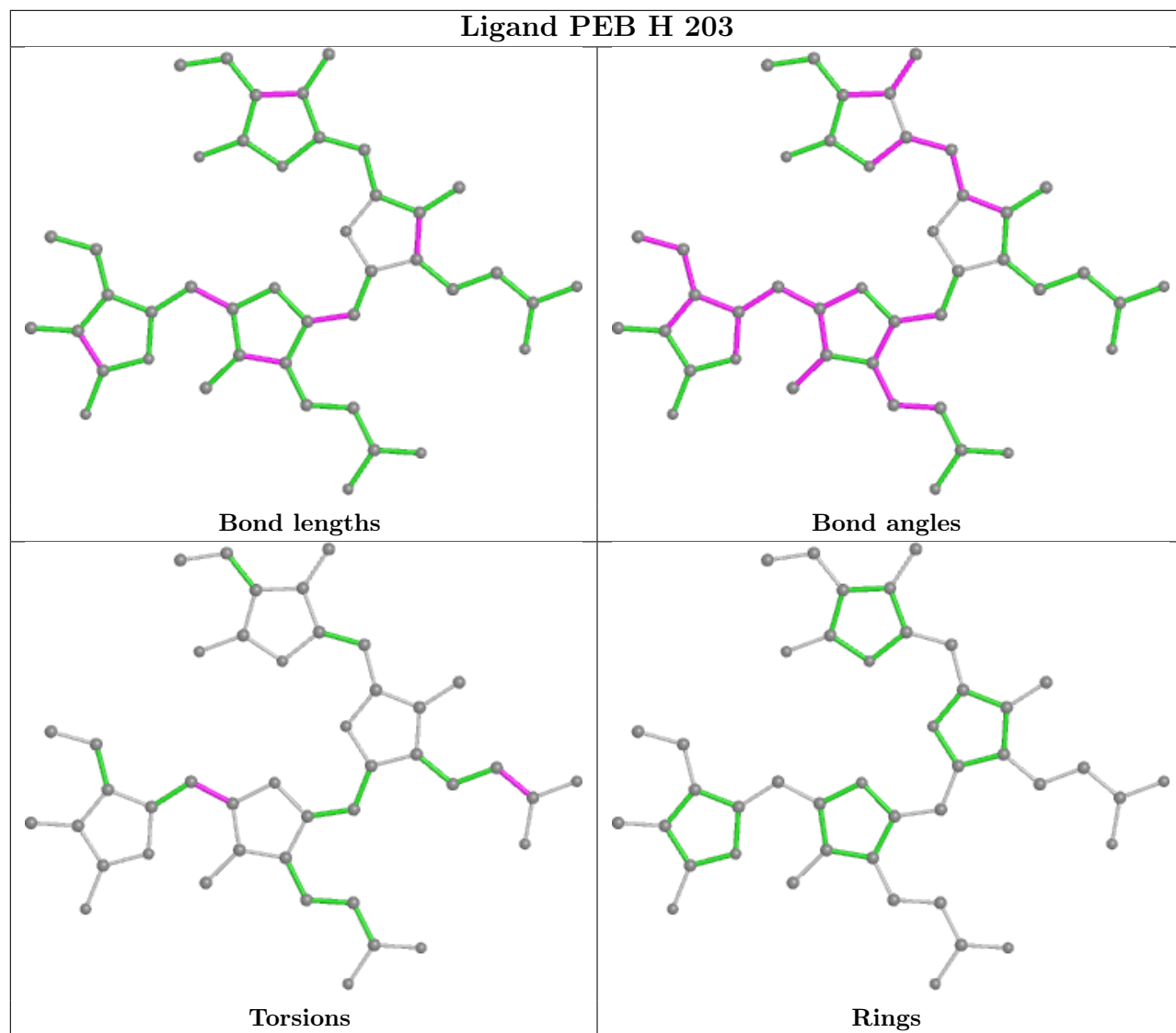


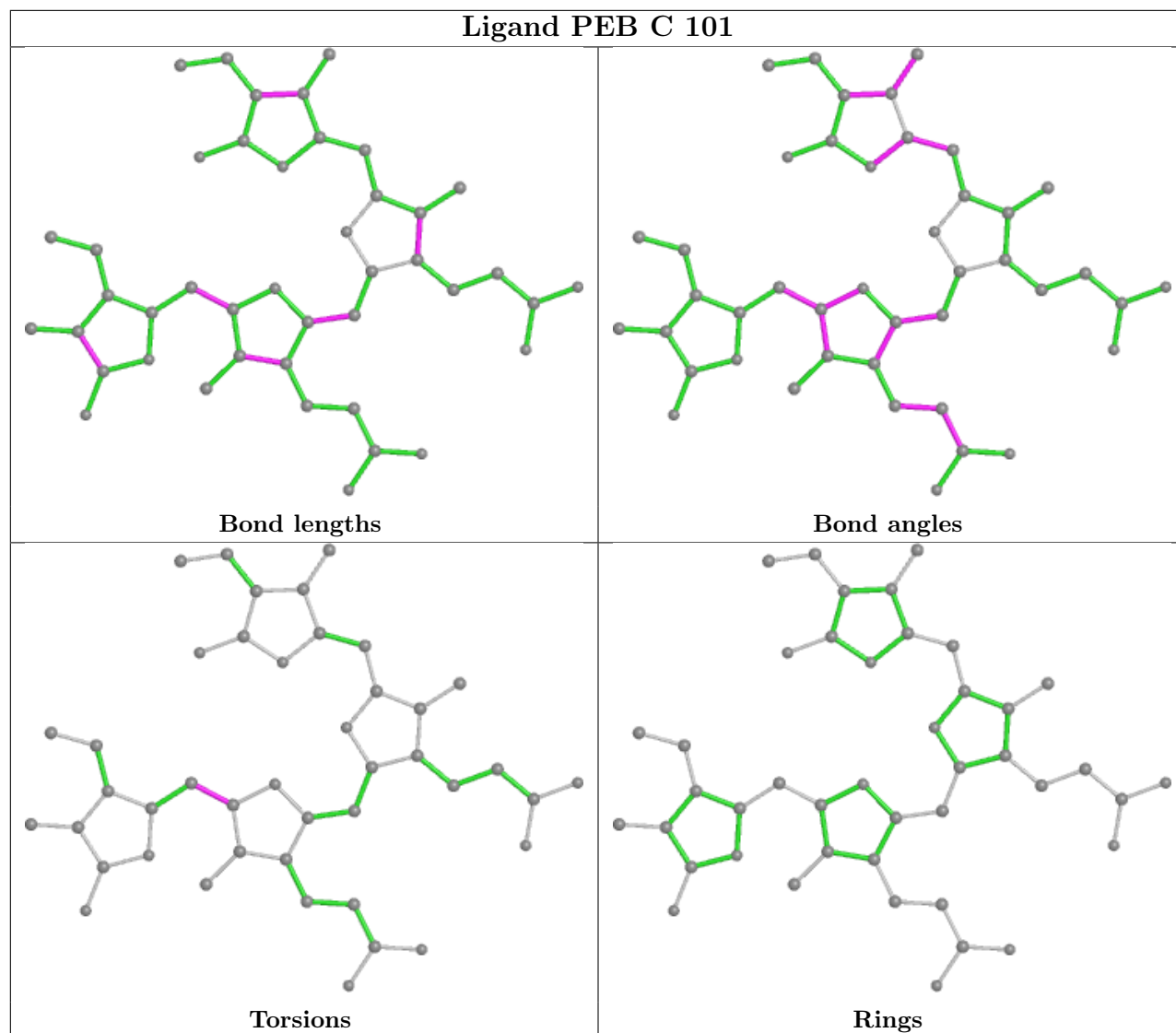




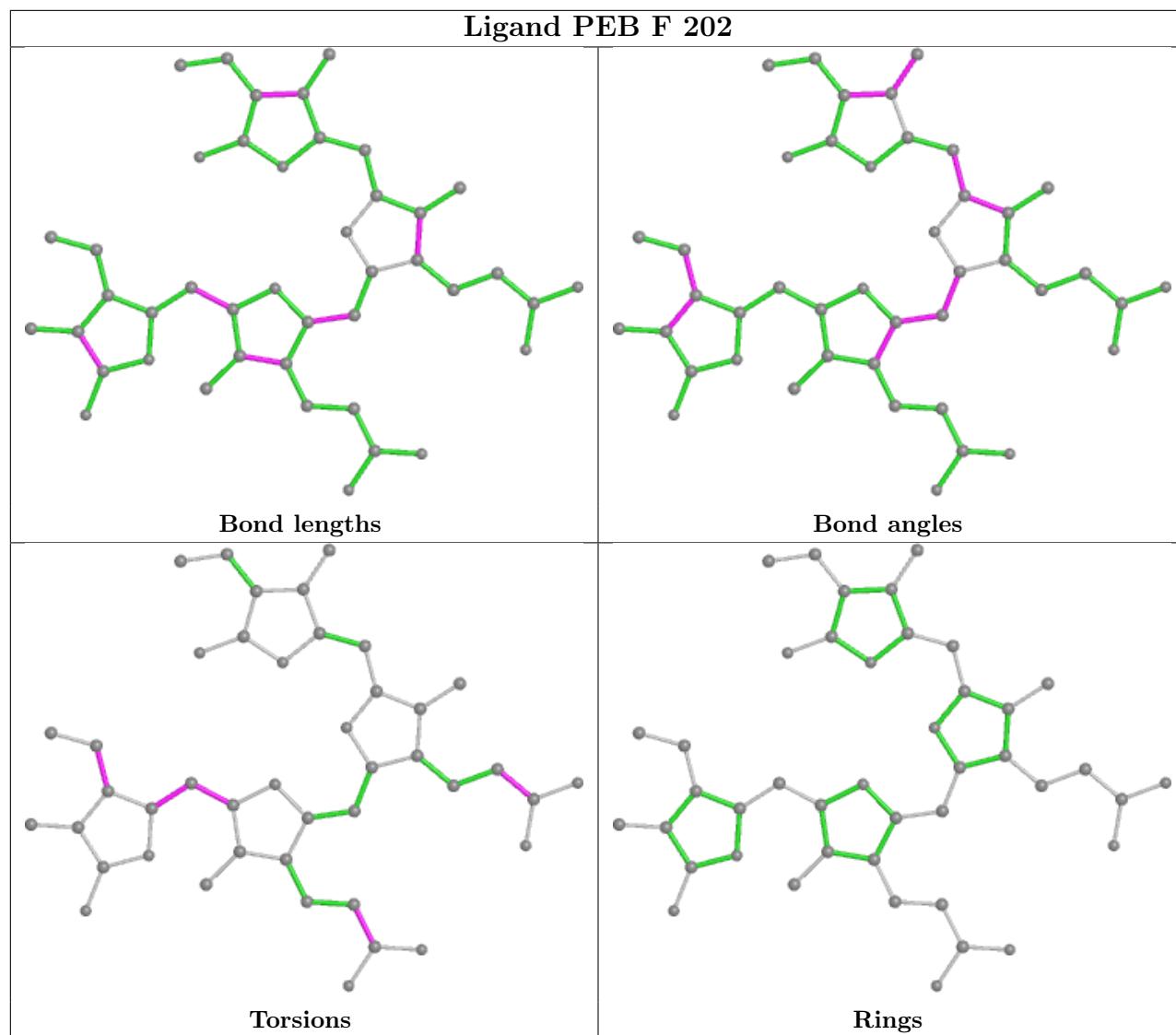


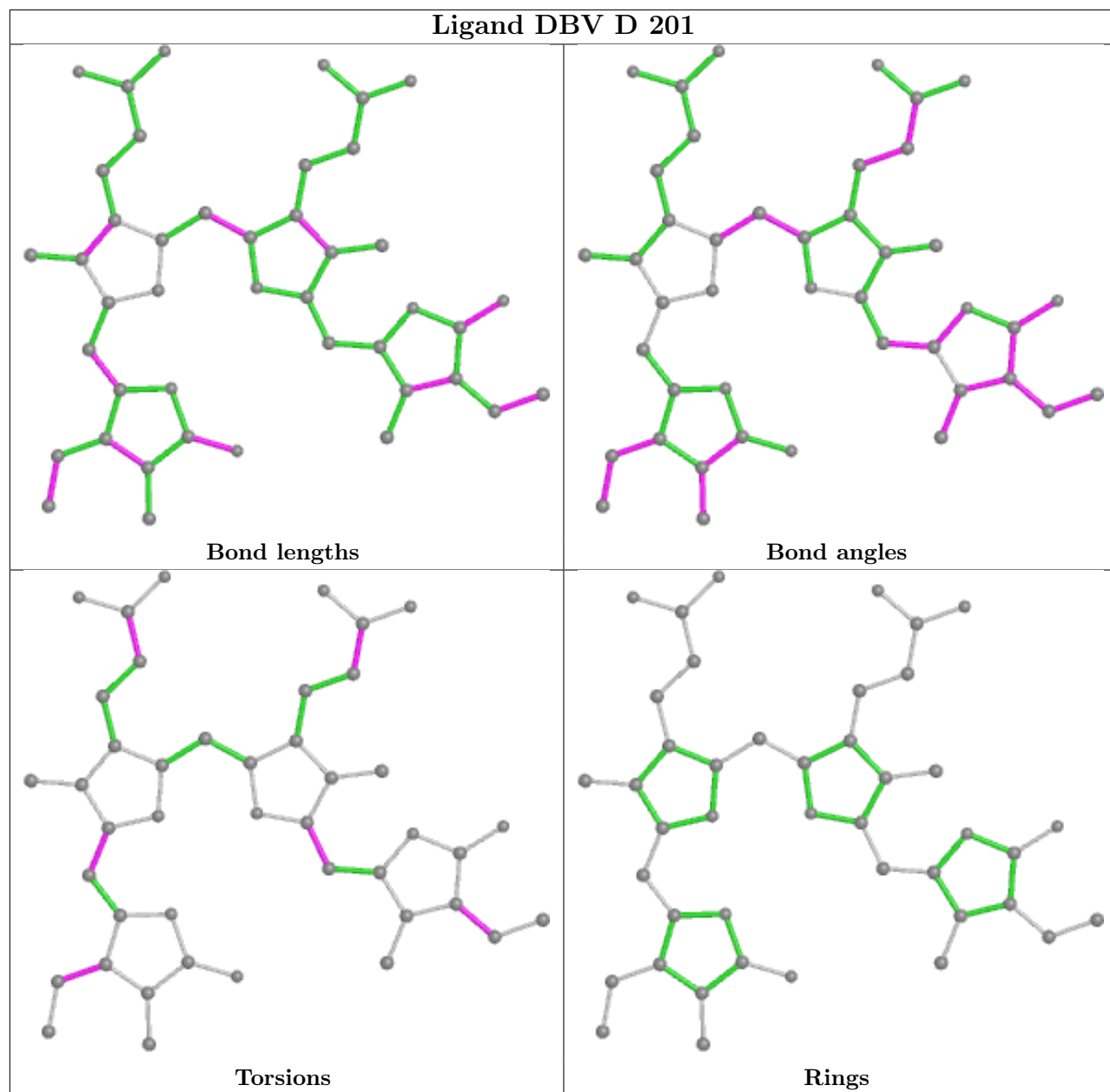


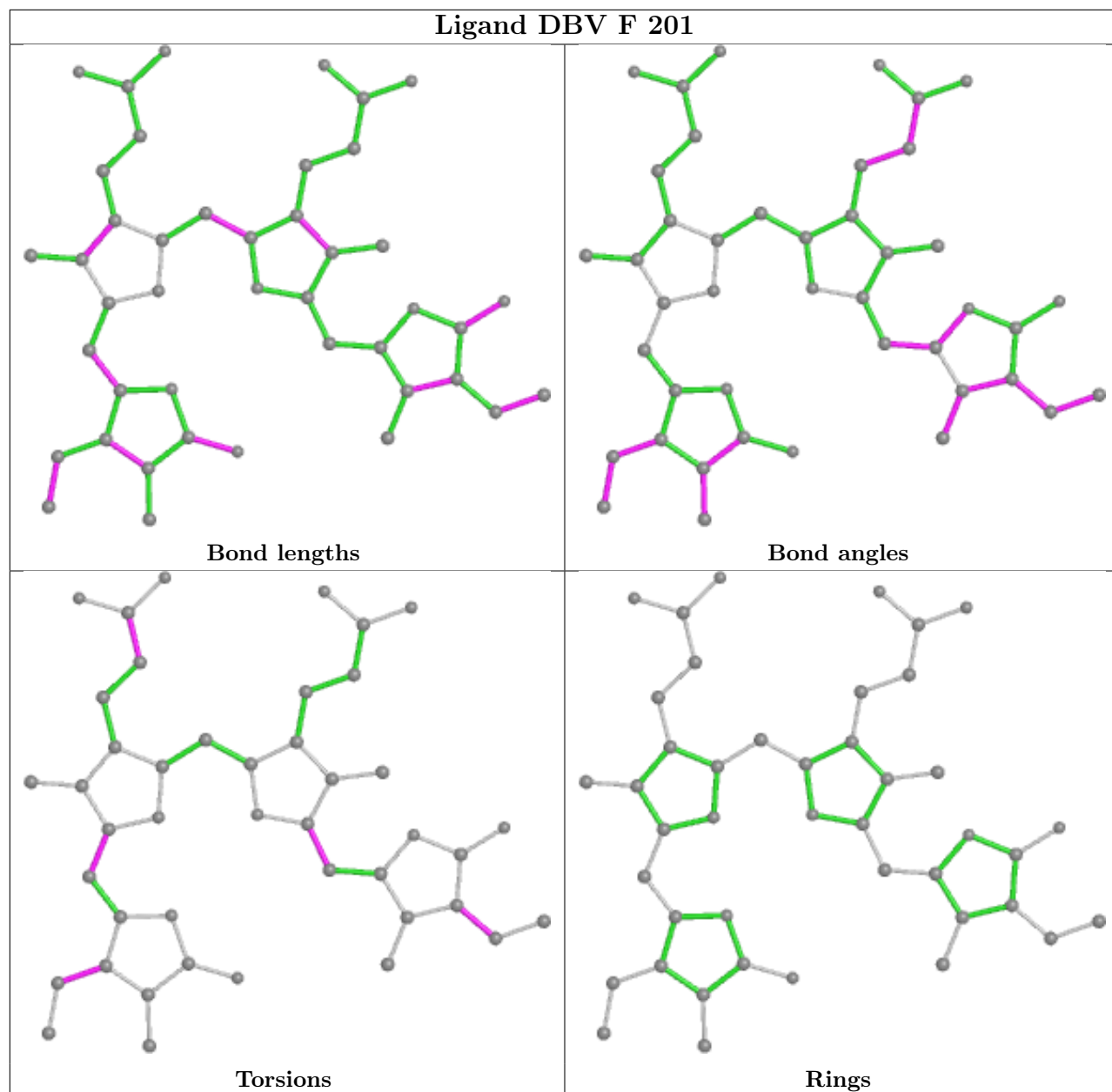


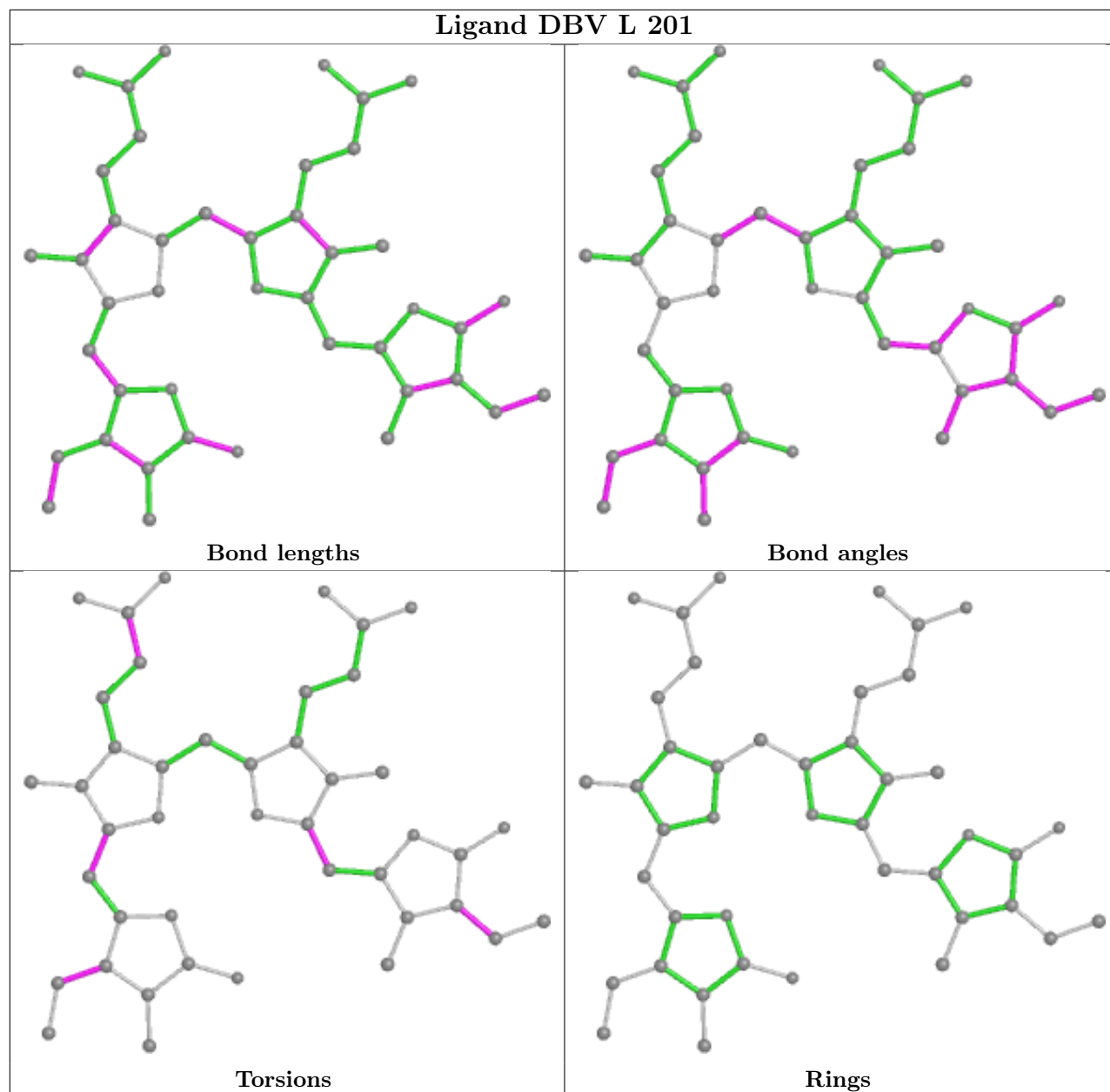


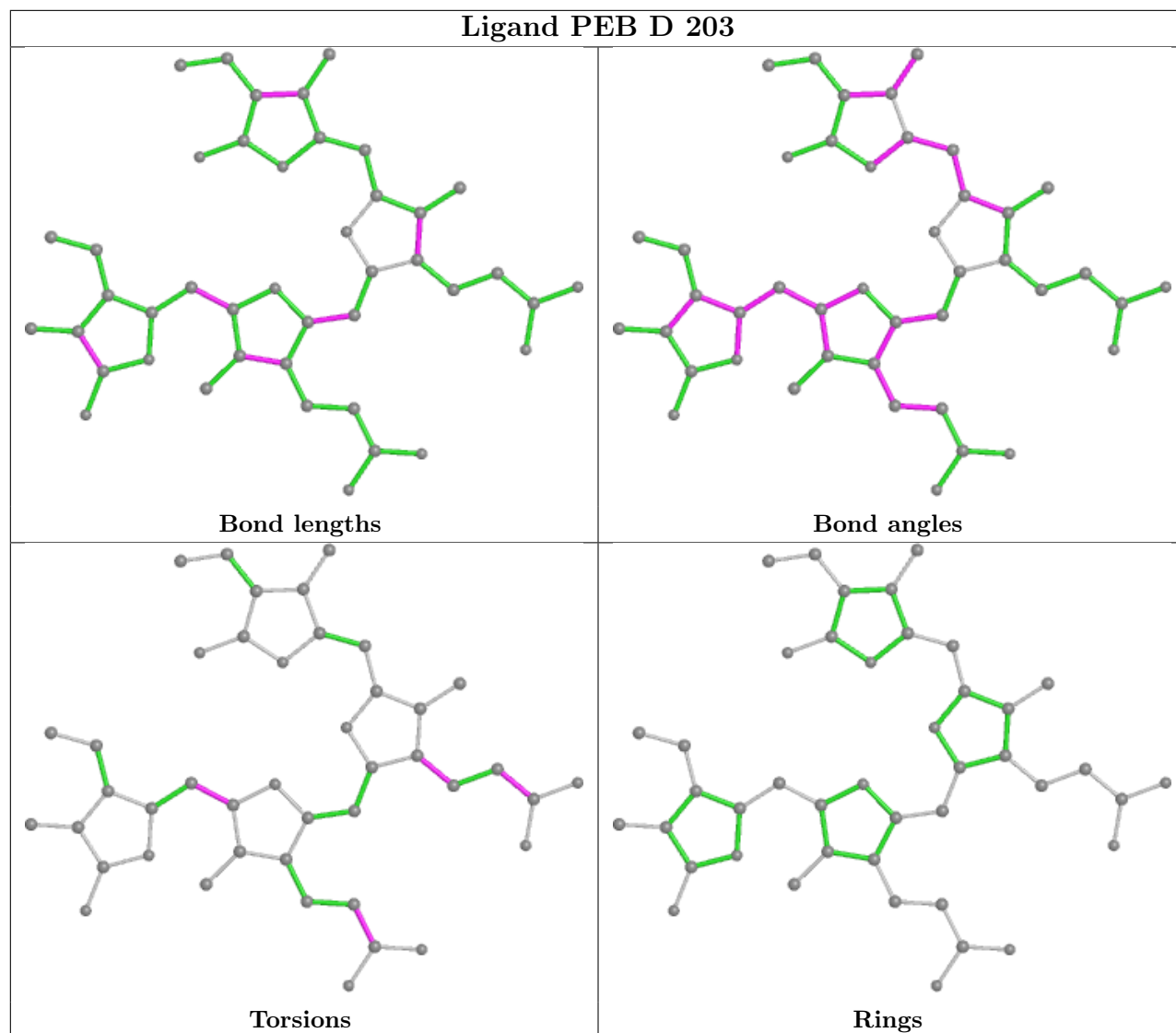


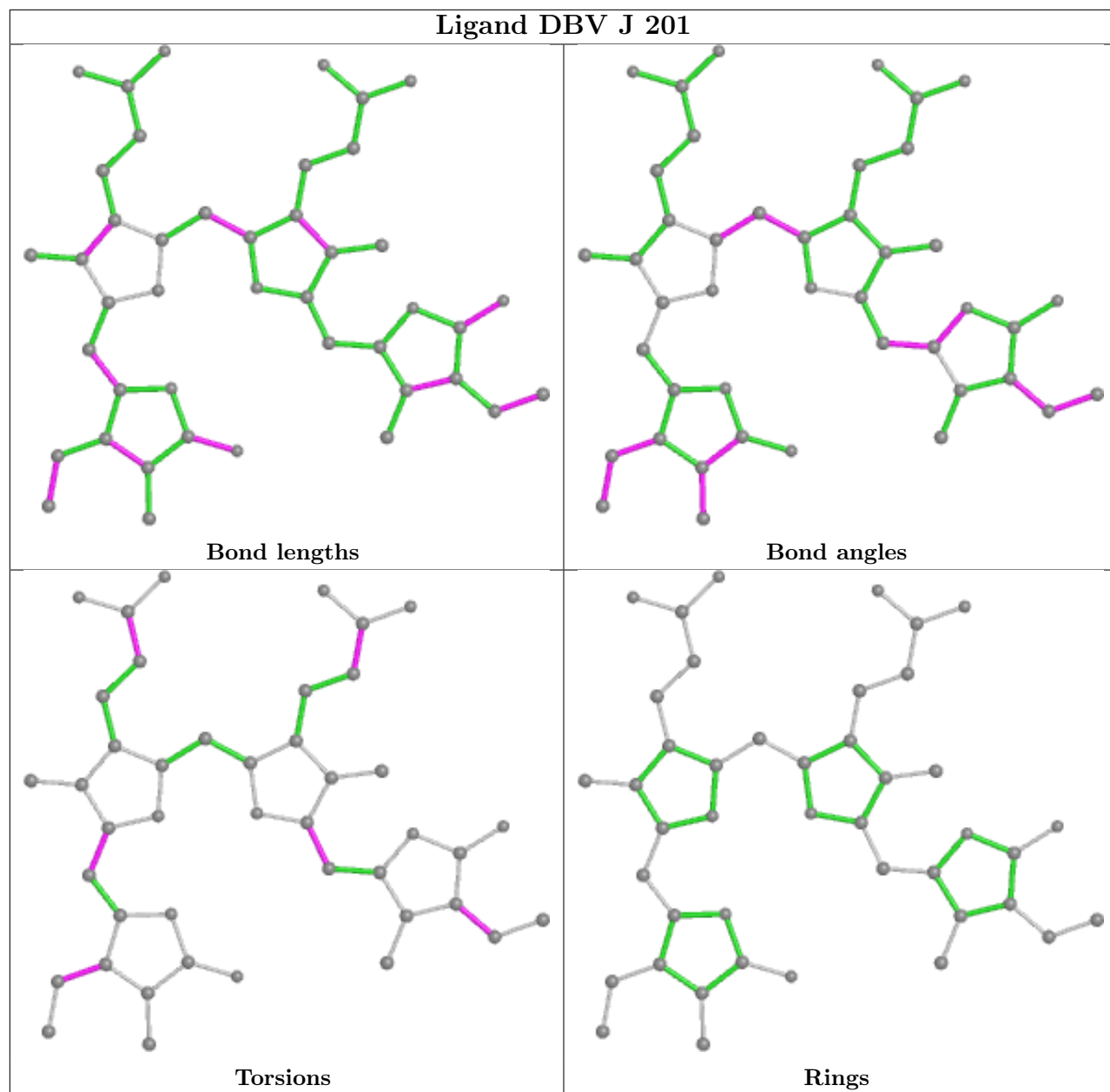


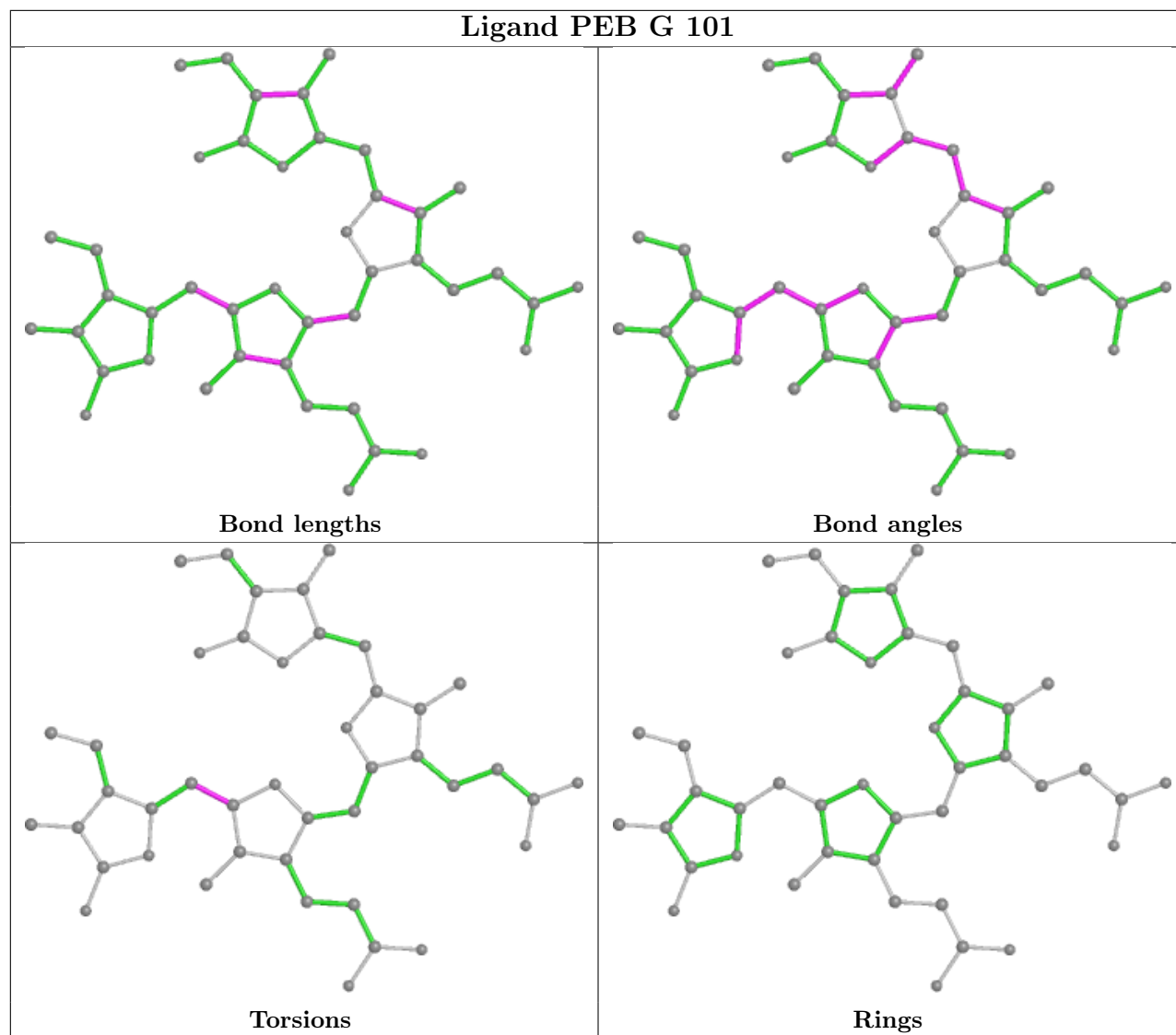


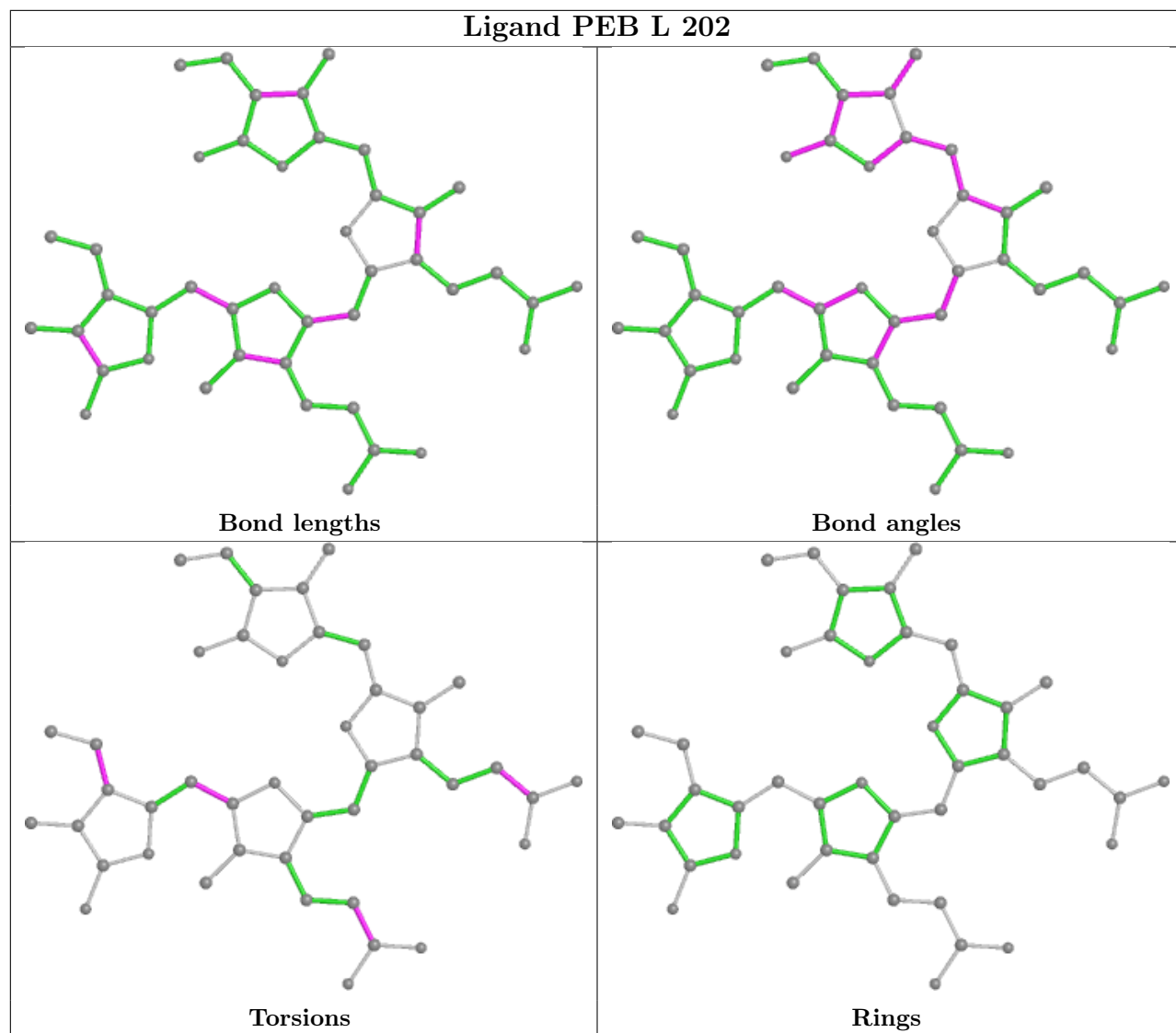




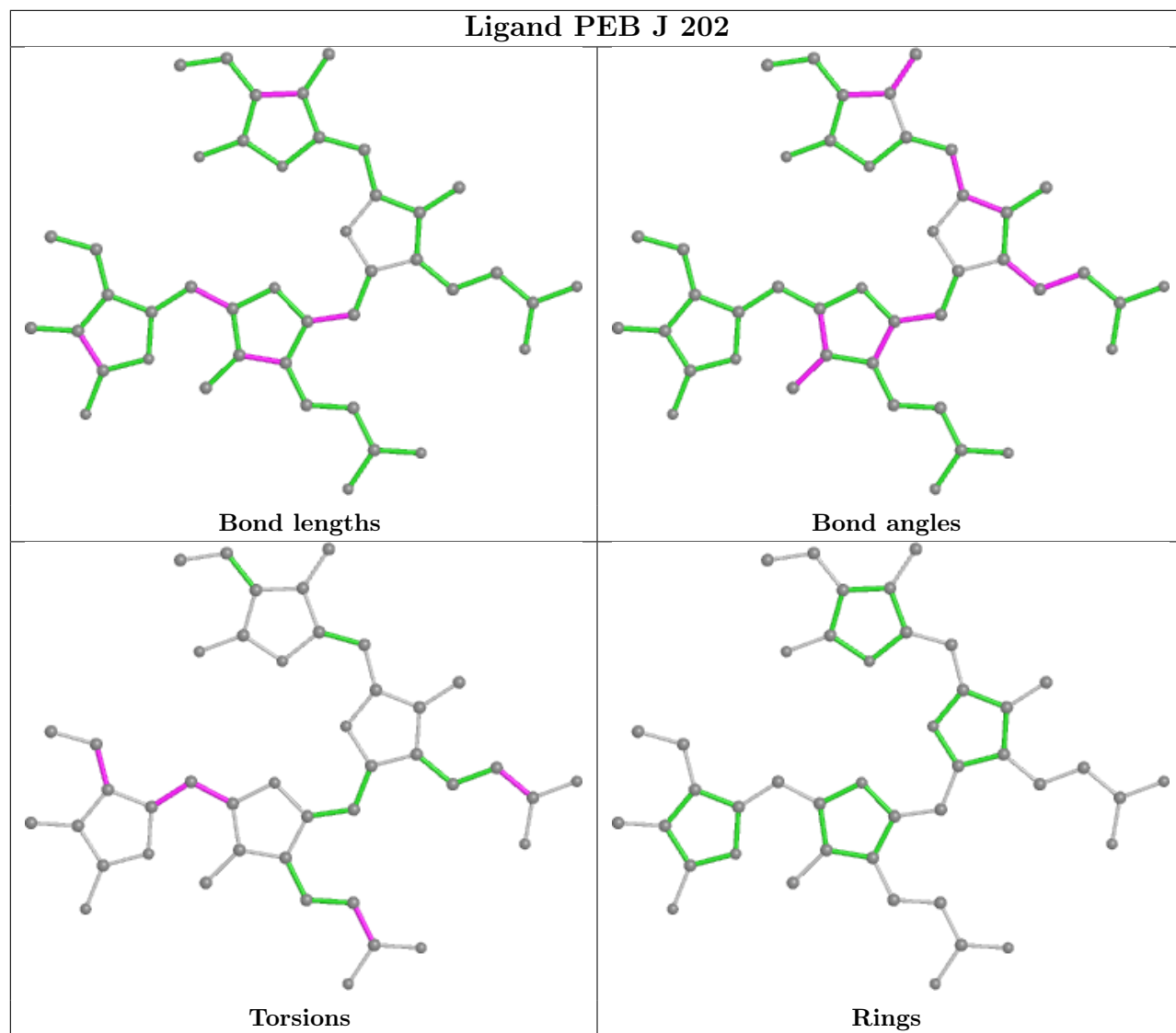


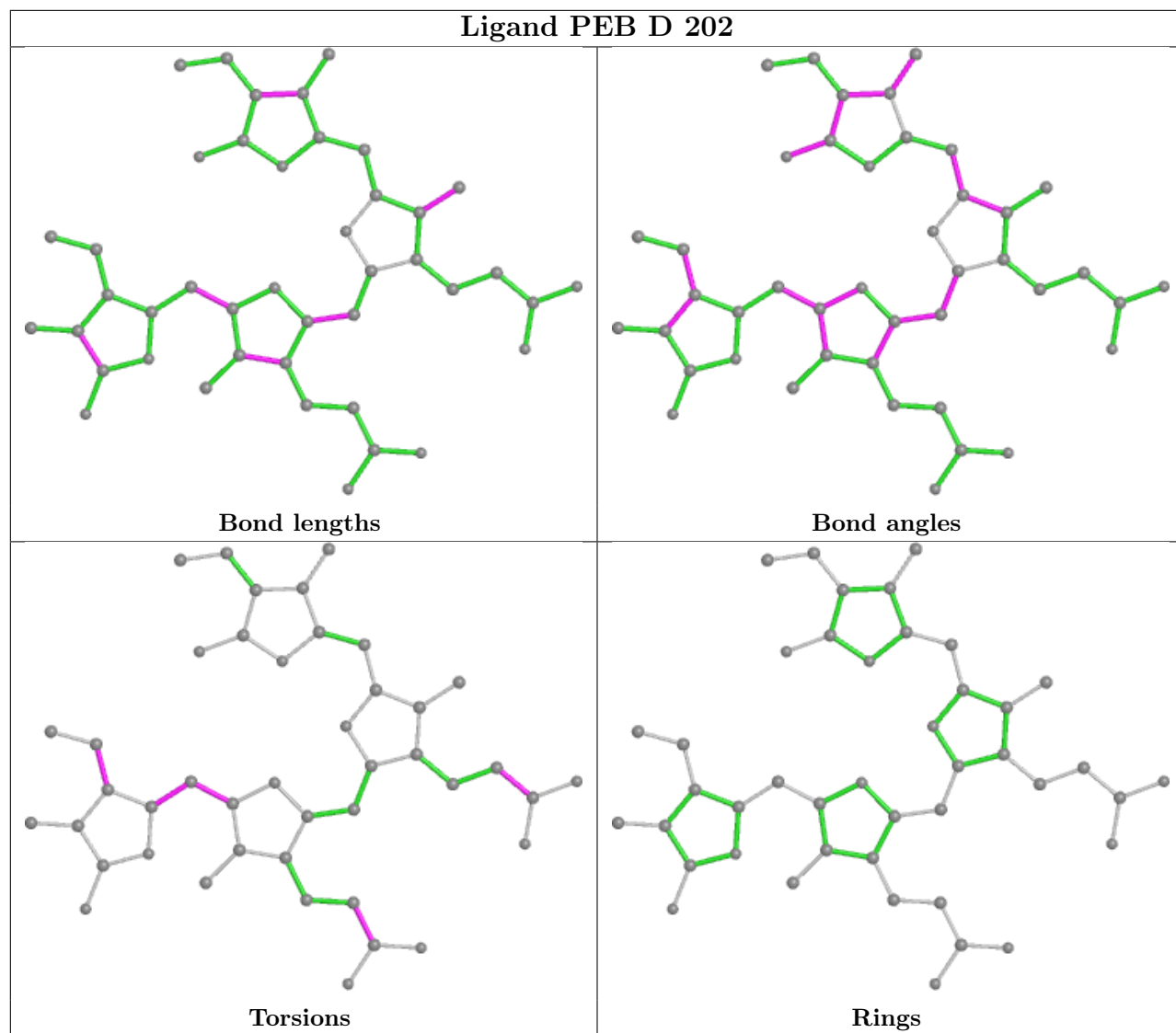


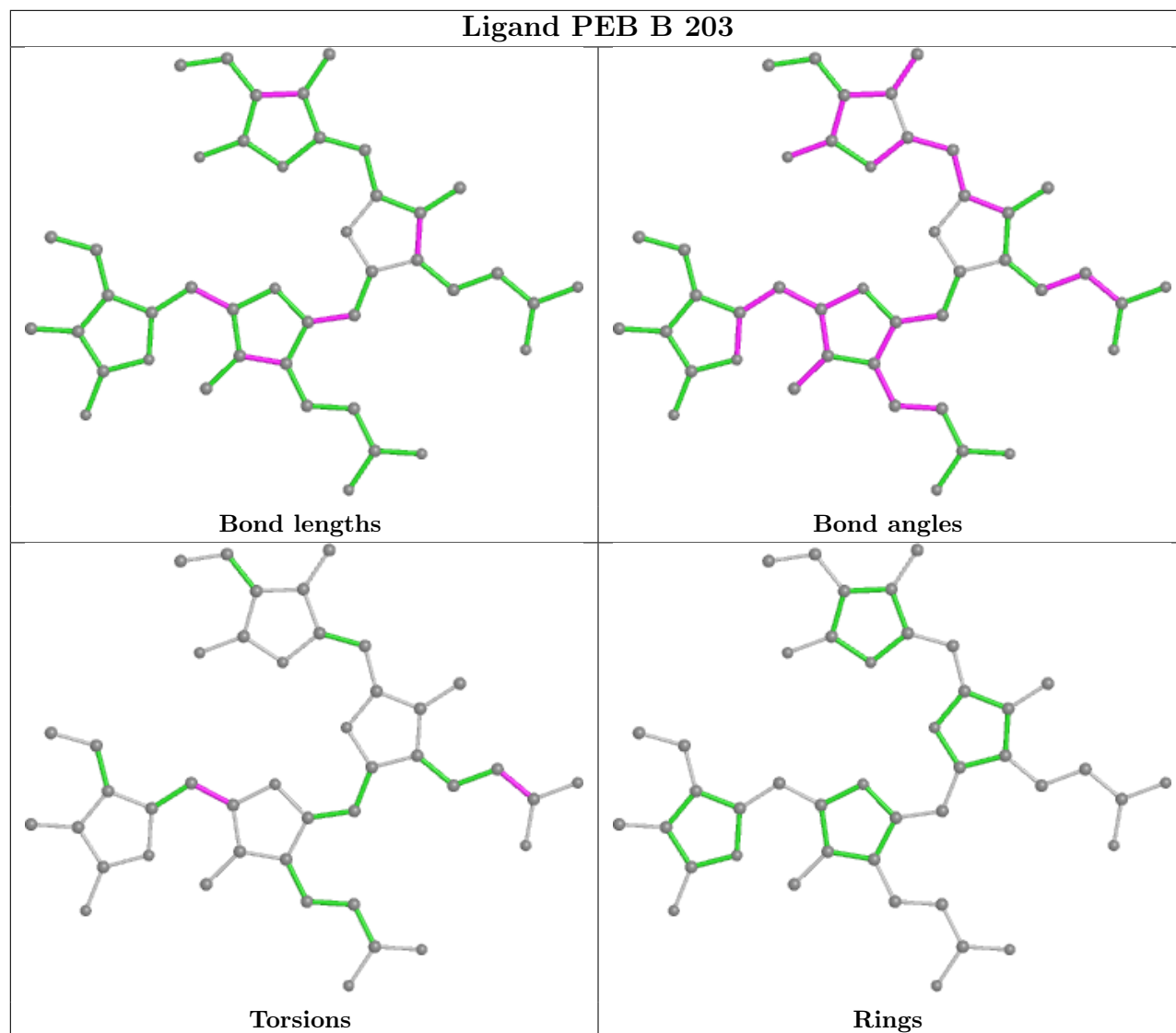


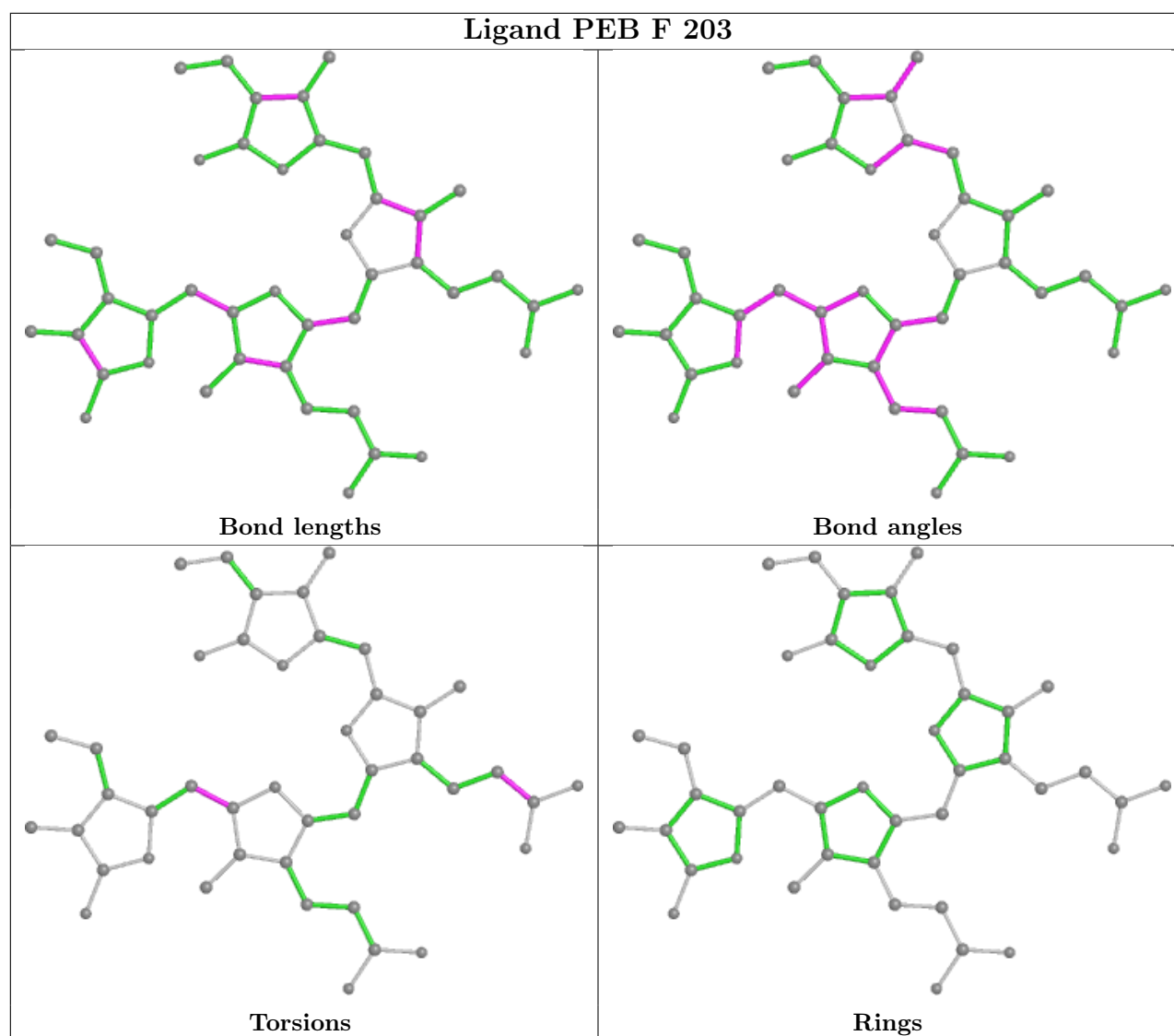












## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
4	I	2
4	G	1
4	E	1
4	K	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	G	62[A]:LEU	C	63:GLY	N	4.49
1	I	62[A]:LEU	C	63:GLY	N	4.44
1	E	62[A]:LEU	C	63:GLY	N	4.24
1	K	62[A]:LEU	C	63:GLY	N	4.05
1	I	58[A]:ASN	C	59[A]:THR	N	1.13

## 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	61/62 (98%)	-0.53	0 <b>100</b> <b>100</b>	10, 19, 35, 44	0
2	B	177/177 (100%)	-0.54	2 (1%) 80 78	9, 18, 41, 58	0
2	D	175/177 (98%)	-0.42	2 (1%) 80 78	10, 19, 45, 91	0
2	F	176/177 (99%)	-0.51	0 <b>100</b> <b>100</b>	9, 18, 40, 51	0
2	H	176/177 (99%)	-0.38	5 (2%) 53 47	10, 19, 59, 98	0
2	J	175/177 (98%)	-0.47	4 (2%) 60 56	10, 18, 40, 80	0
2	L	175/177 (98%)	-0.10	7 (4%) 38 32	12, 27, 53, 100	0
3	C	65/67 (97%)	-0.47	1 (1%) 73 70	11, 18, 35, 51	0
4	E	72/74 (97%)	-0.41	0 <b>100</b> <b>100</b>	10, 19, 33, 38	18 (25%)
4	G	71/74 (95%)	-0.24	1 (1%) 75 72	11, 21, 39, 43	17 (23%)
4	I	72/74 (97%)	-0.40	0 <b>100</b> <b>100</b>	12, 20, 36, 43	18 (25%)
4	K	72/74 (97%)	-0.25	2 (2%) 53 47	14, 23, 37, 48	18 (25%)
All	All	1467/1487 (98%)	-0.40	24 (1%) 68 68	9, 20, 42, 100	71 (4%)

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	3	ASP	6.2
2	L	3	ASP	5.3
2	L	30	PHE	4.2
2	H	12	ALA	4.2
2	L	12	ALA	3.7

### 6.2 Non-standard residues in protein, DNA, RNA chains [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
4	LYZ	K	4	10/11	0.90	0.16	30,35,45,49	0
1	LYZ	A	4	10/11	0.94	0.14	22,26,43,49	0
4	LYZ	E	4	10/11	0.95	0.10	20,25,35,42	0
4	LYZ	I	4	10/11	0.96	0.10	18,22,34,41	0
3	LYZ	C	4	10/11	0.96	0.09	19,22,39,40	0
4	LYZ	G	4	10/11	0.97	0.07	17,21,34,40	0

### 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.

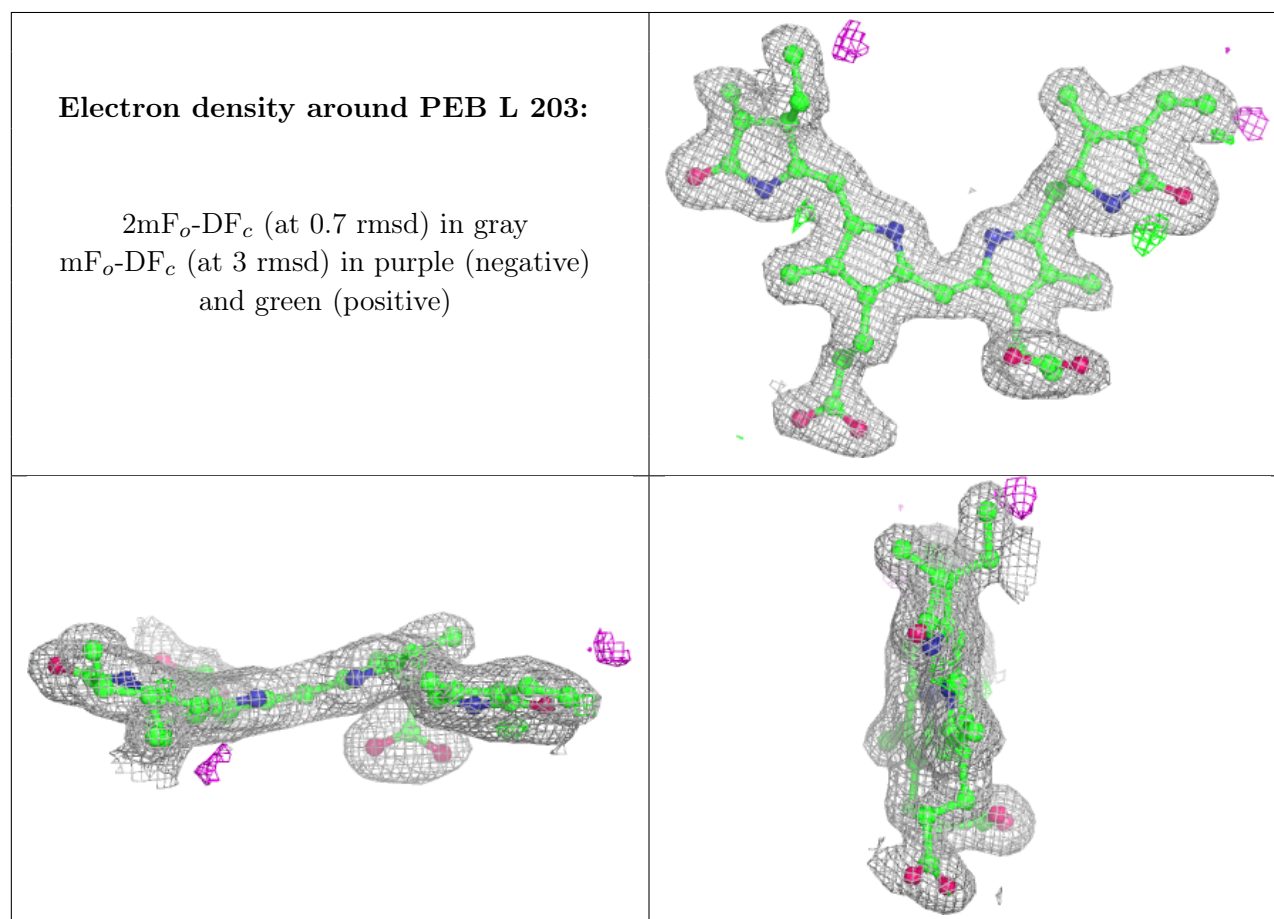
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	PEB	L	203	43/43	0.93	0.11	11,19,34,45	0
6	DBV	L	201	43/43	0.93	0.09	16,22,32,42	0
5	PEB	G	101	43/43	0.94	0.09	14,22,51,69	0
5	PEB	H	202	43/43	0.94	0.09	10,17,27,30	0
5	PEB	I	101	43/43	0.94	0.10	12,18,39,67	0
5	PEB	K	101	43/43	0.94	0.09	11,16,41,56	0
5	PEB	A	101	43/43	0.94	0.10	8,16,41,57	0
5	PEB	C	101	43/43	0.94	0.10	8,14,35,48	0
5	PEB	E	101	43/43	0.95	0.09	6,13,26,50	0
5	PEB	J	202	43/43	0.95	0.08	9,14,28,41	0
5	PEB	J	203	43/43	0.95	0.09	8,13,24,35	0
5	PEB	F	203	43/43	0.95	0.08	9,16,33,39	0
5	PEB	L	202	43/43	0.95	0.08	14,20,29,32	0
5	PEB	D	202	43/43	0.95	0.08	6,14,22,29	0
6	DBV	B	201	43/43	0.95	0.08	10,16,28,42	0
6	DBV	D	201	43/43	0.95	0.09	11,16,32,36	0
6	DBV	F	201	43/43	0.95	0.08	11,18,32,44	0
6	DBV	H	201	43/43	0.95	0.08	12,20,28,39	0
6	DBV	J	201	43/43	0.95	0.09	10,16,26,41	0
5	PEB	D	203	43/43	0.95	0.09	7,13,24,32	0

*Continued on next page...*

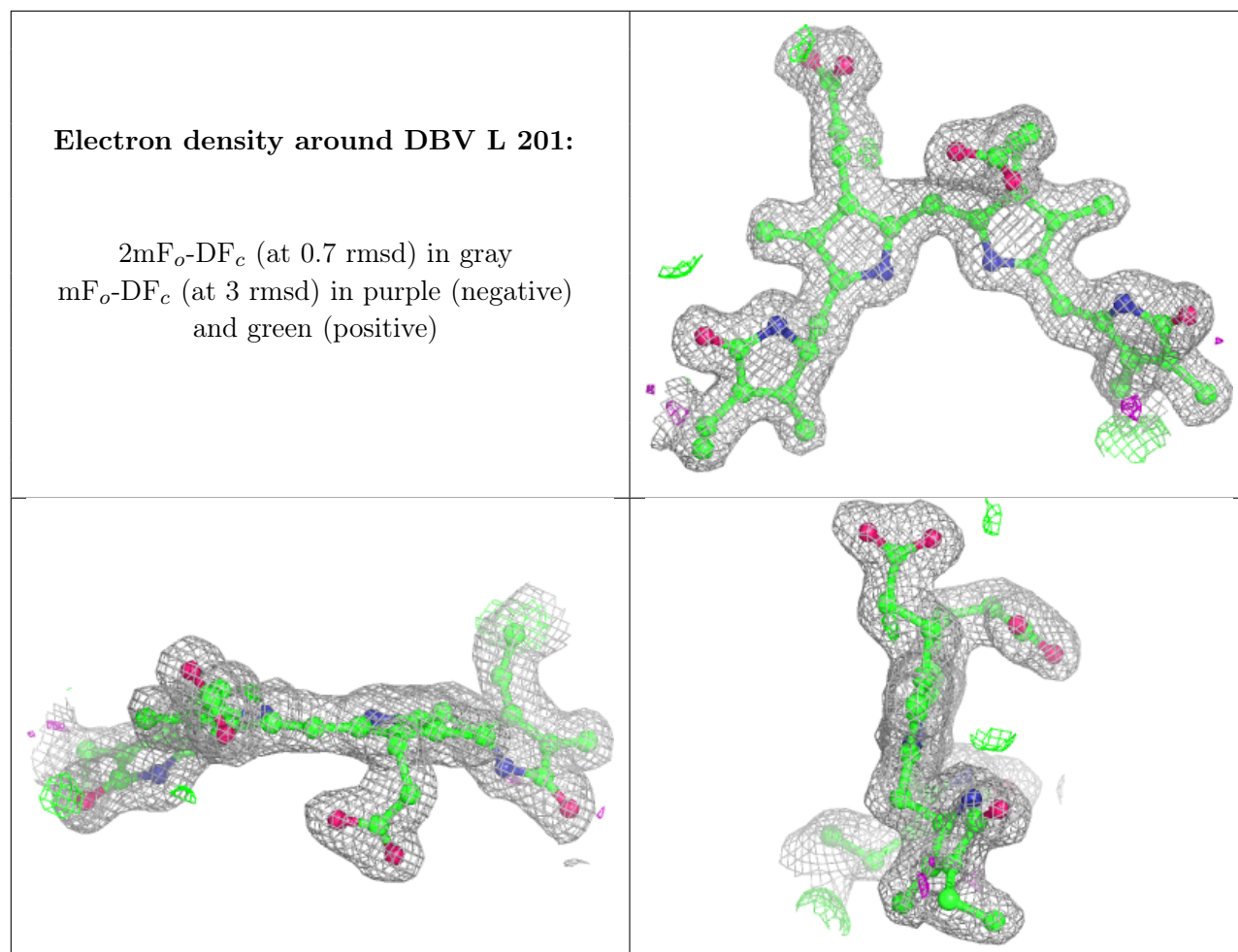
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	PEB	B	202	43/43	0.96	0.08	8,15,29,32	0
5	PEB	B	203	43/43	0.96	0.08	7,13,21,38	0
5	PEB	F	202	43/43	0.96	0.08	7,14,22,27	0
5	PEB	H	203	43/43	0.97	0.08	8,12,19,37	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.

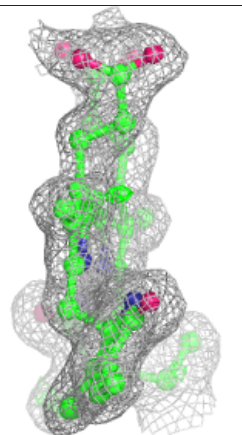
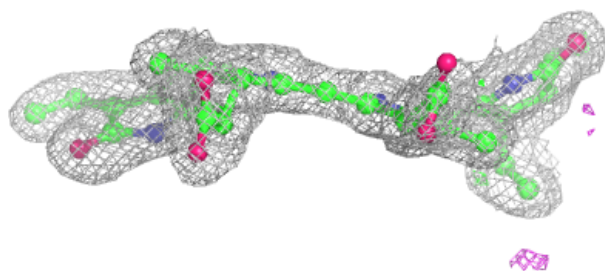
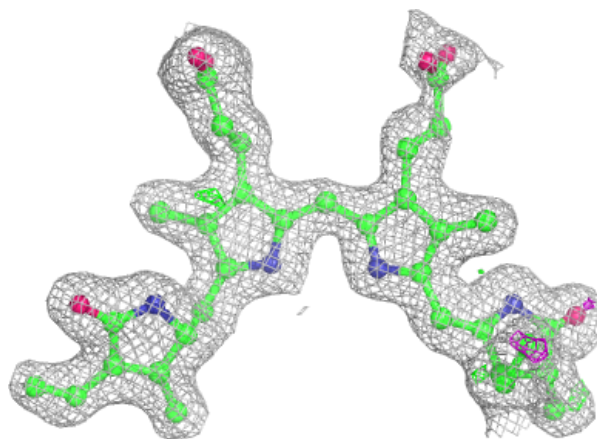




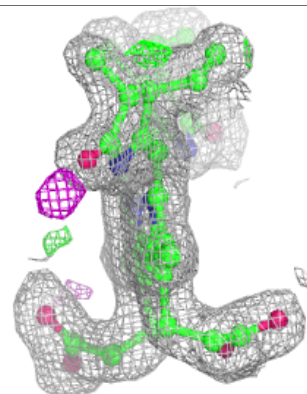
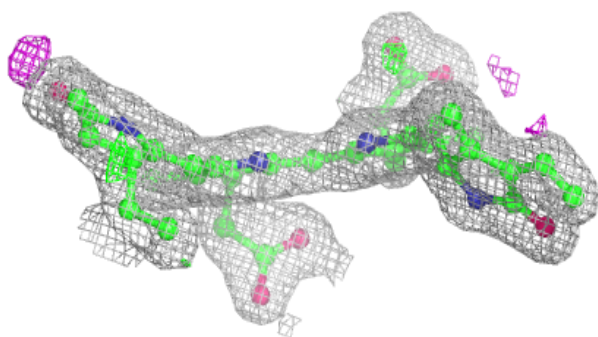
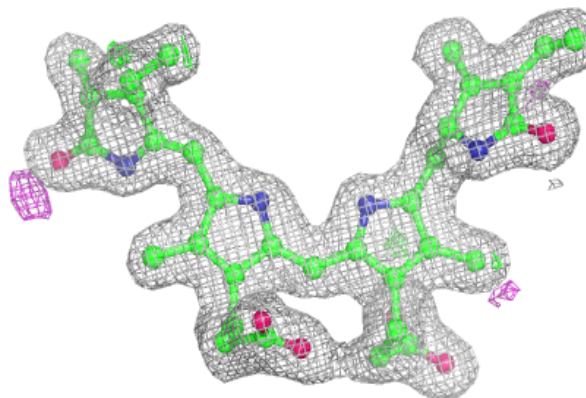


**Electron density around PEB G 101:**

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

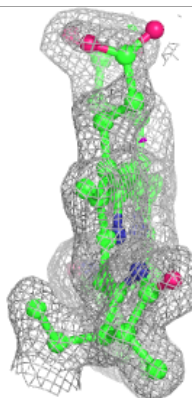
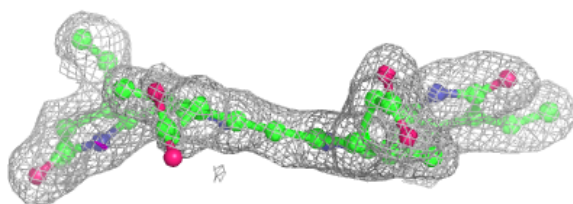
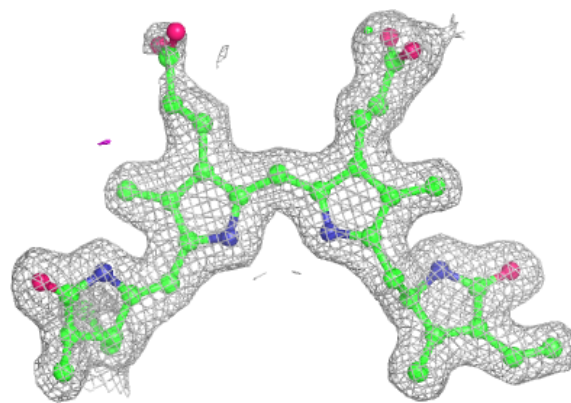
**Electron density around PEB H 202:**

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

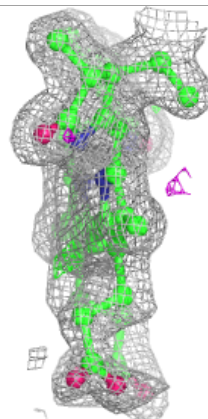
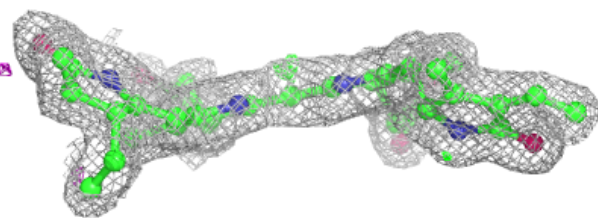
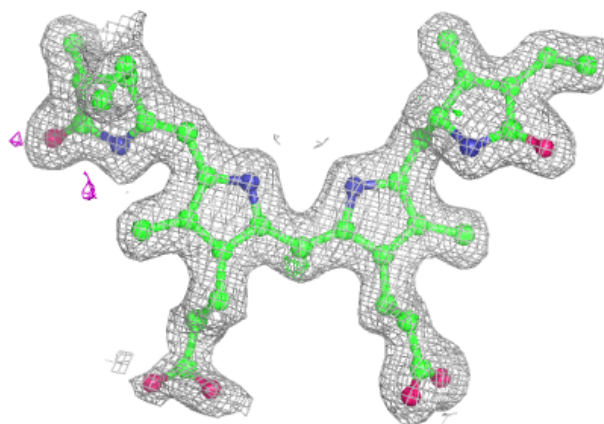


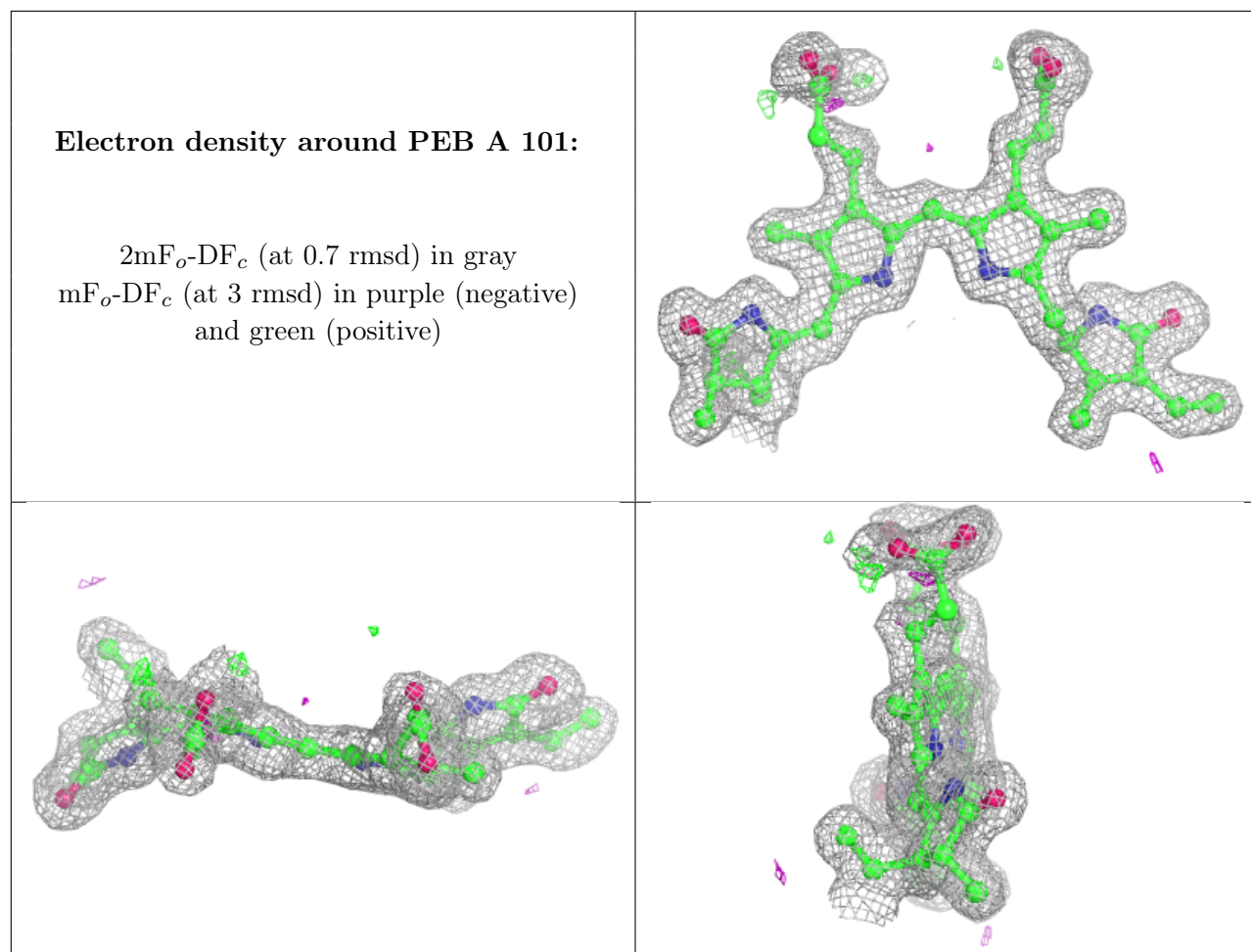
**Electron density around PEB I 101:**

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

**Electron density around PEB K 101:**

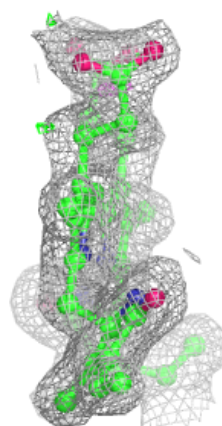
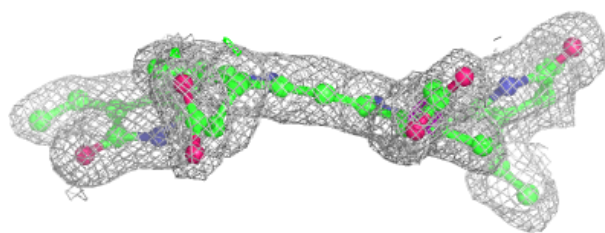
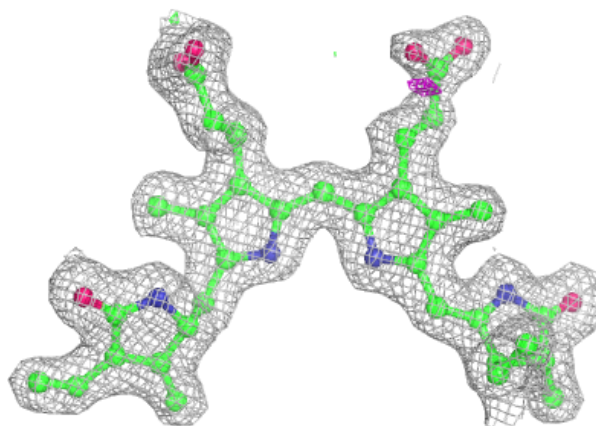
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





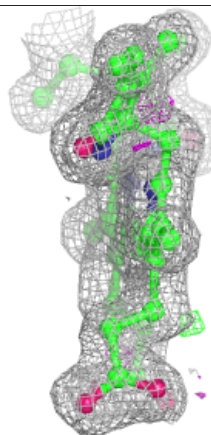
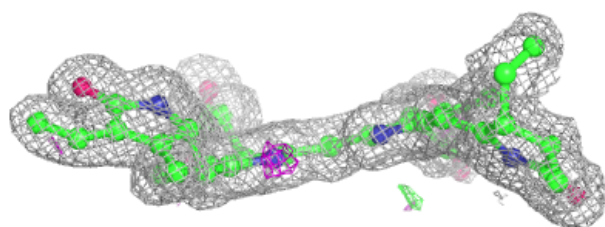
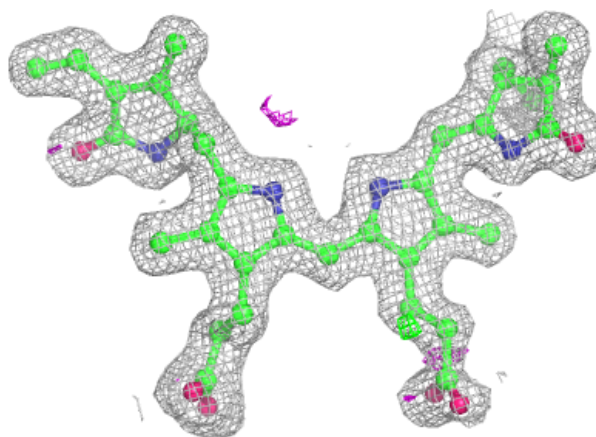
**Electron density around PEB C 101:**

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

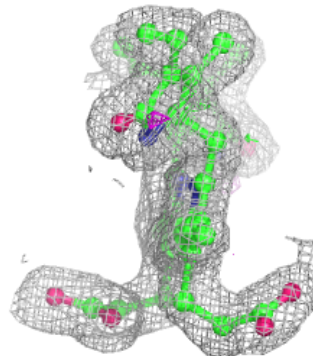
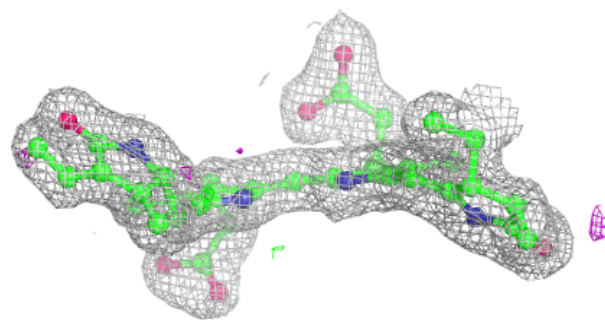
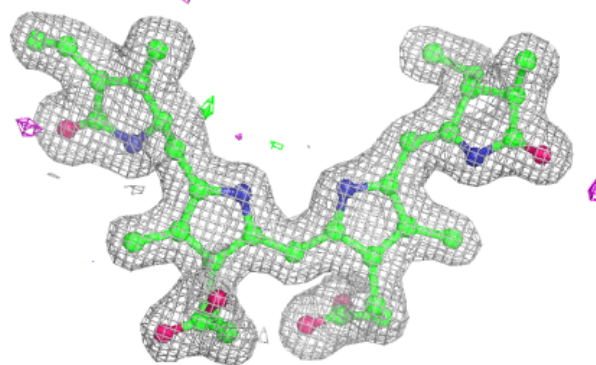


**Electron density around PEB E 101:**

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

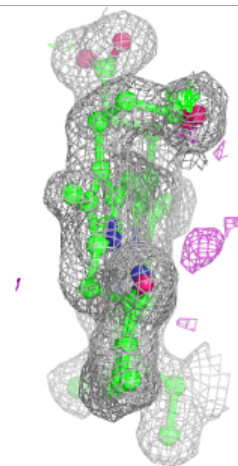
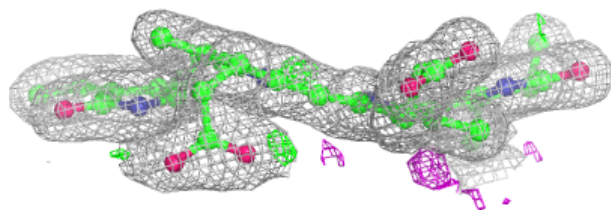
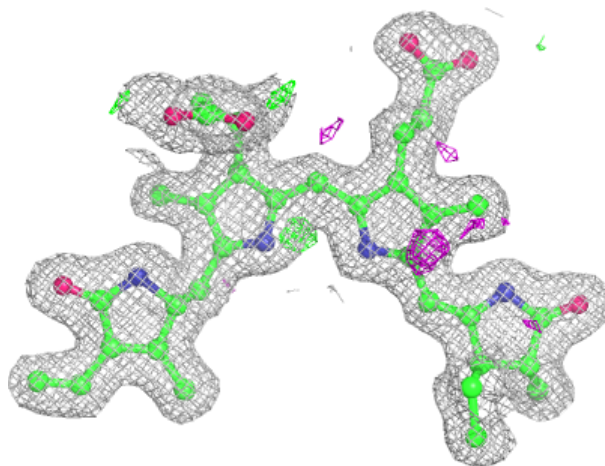
**Electron density around PEB J 202:**

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



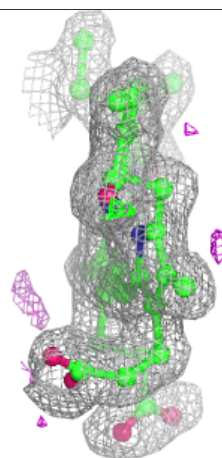
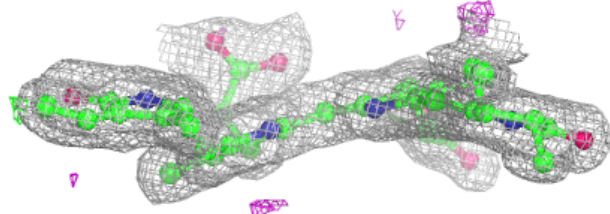
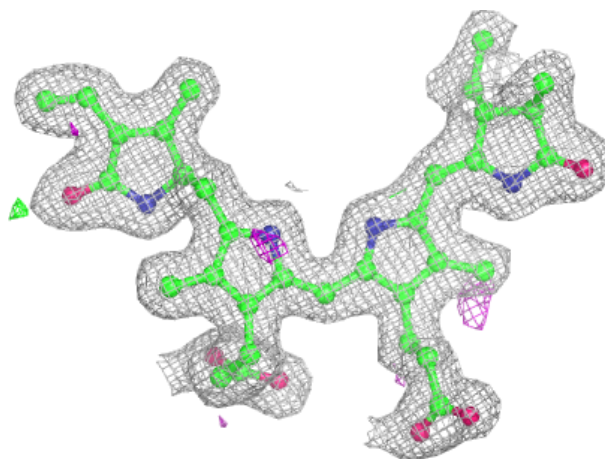
**Electron density around PEB J 203:**

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



**Electron density around PEB F 203:**

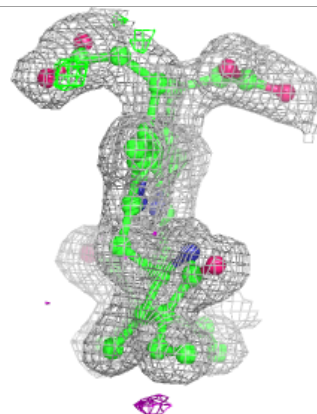
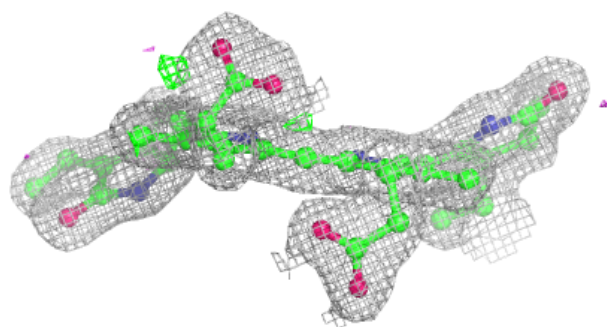
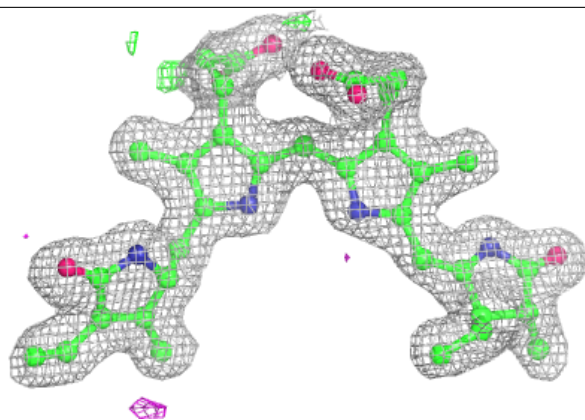
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



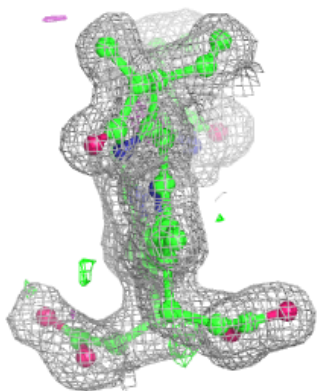
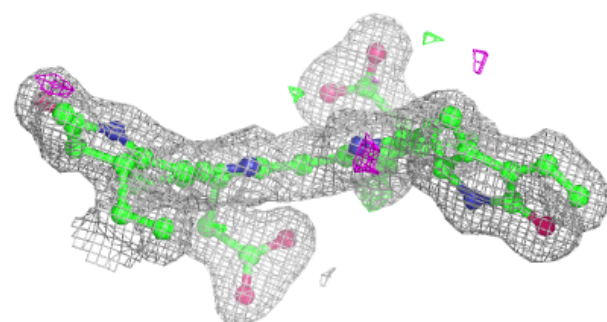
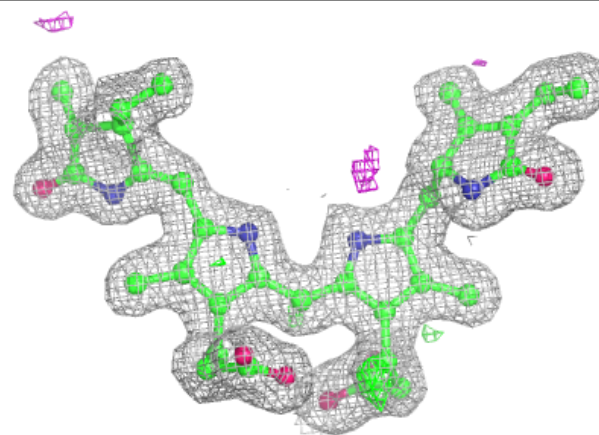


**Electron density around PEB L 202:**

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

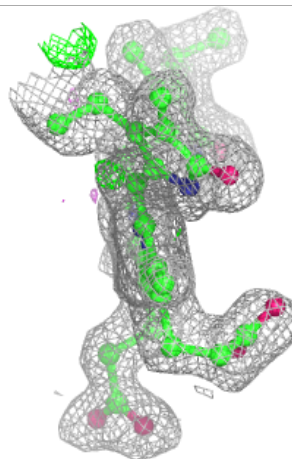
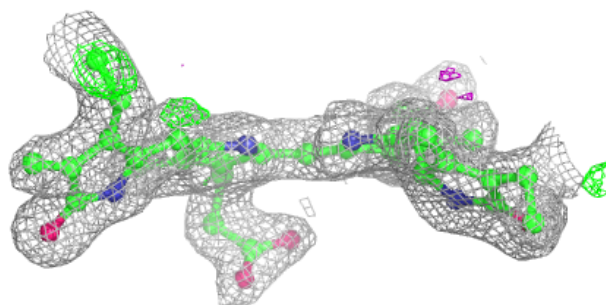
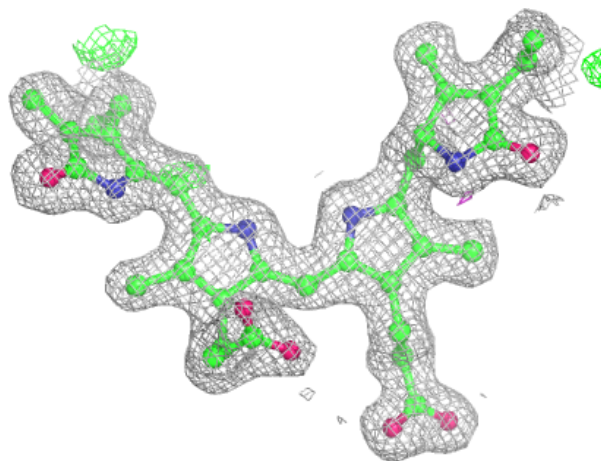
**Electron density around PEB D 202:**

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



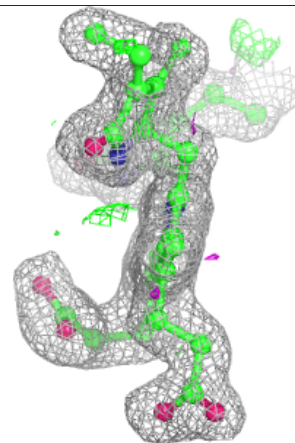
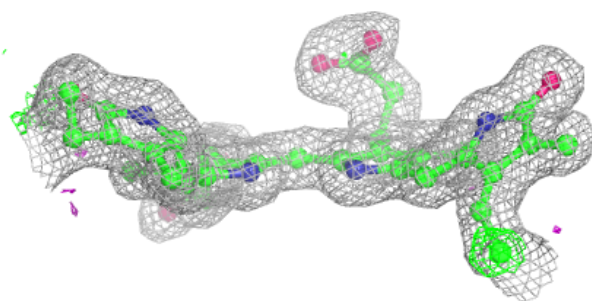
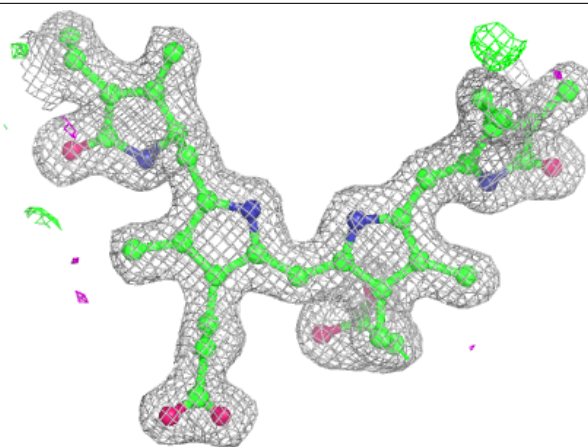
**Electron density around DBV B 201:**

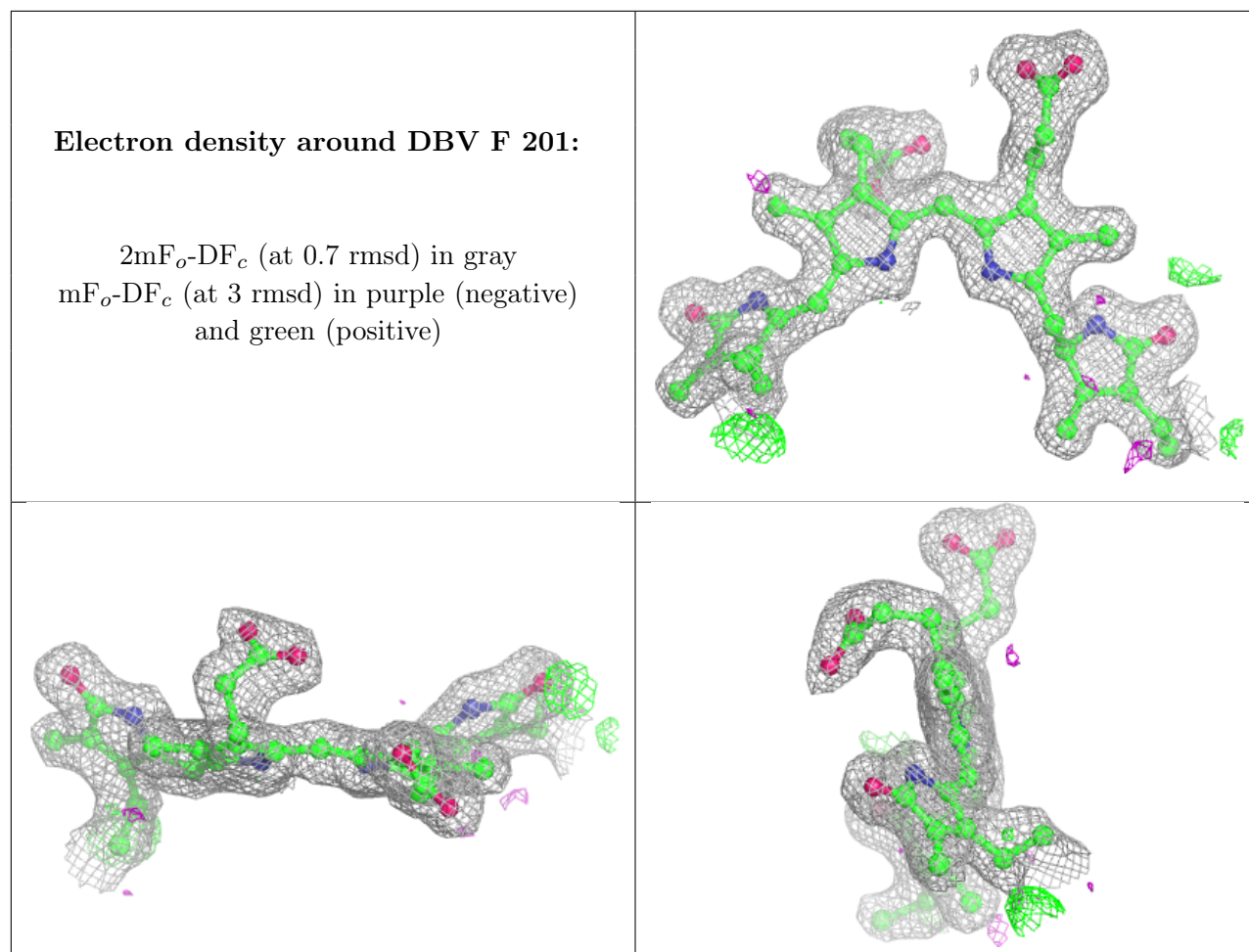
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around DBV D 201:**

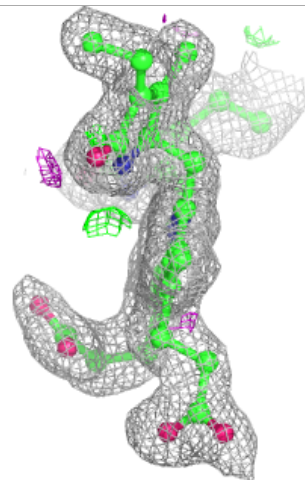
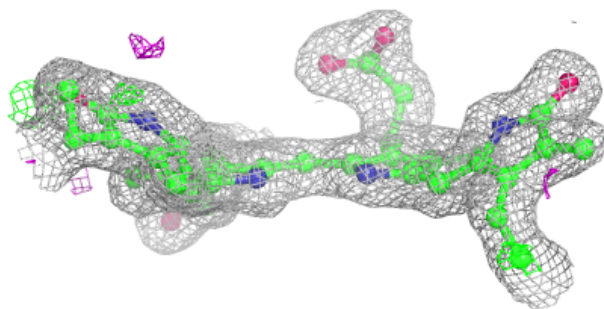
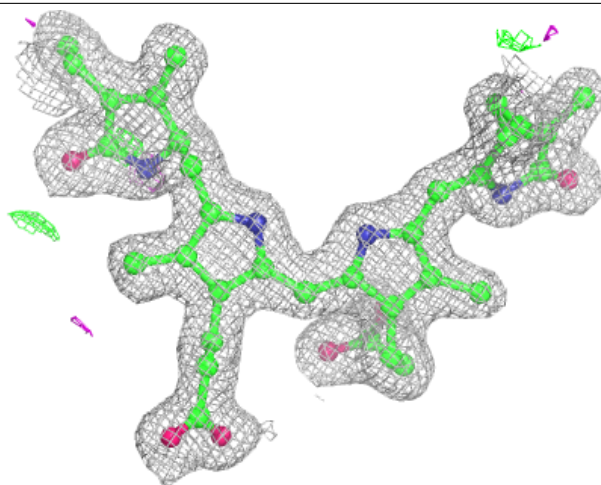
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





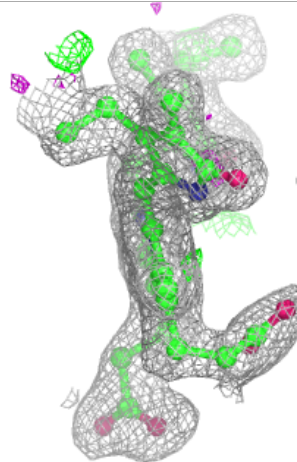
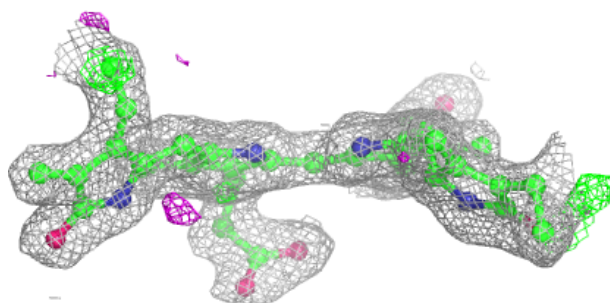
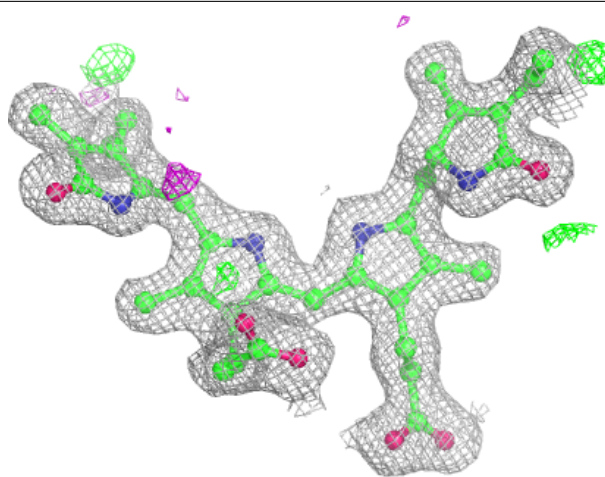
**Electron density around DBV H 201:**

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



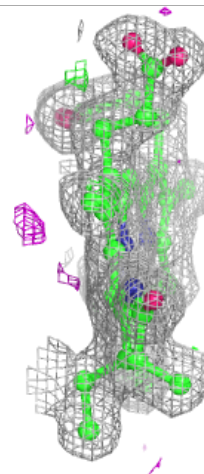
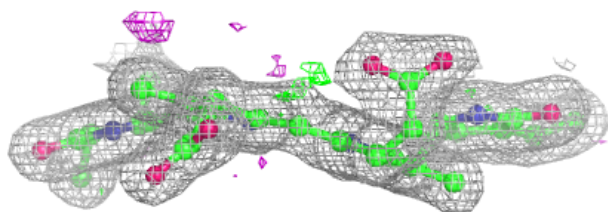
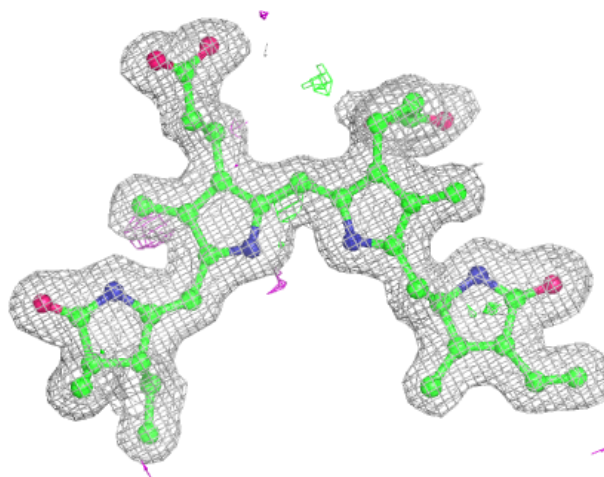
**Electron density around DBV J 201:**

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



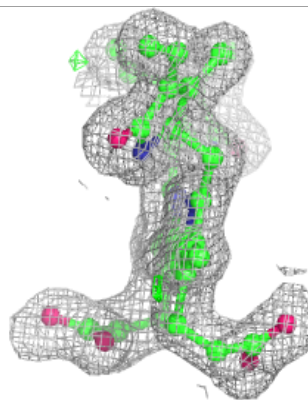
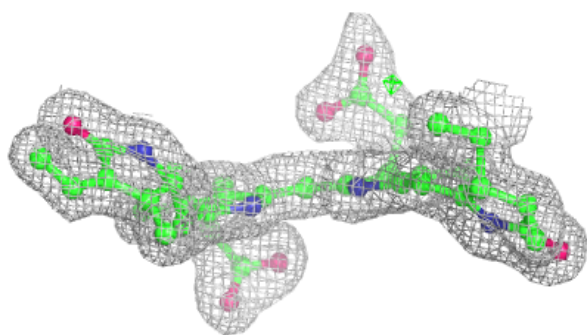
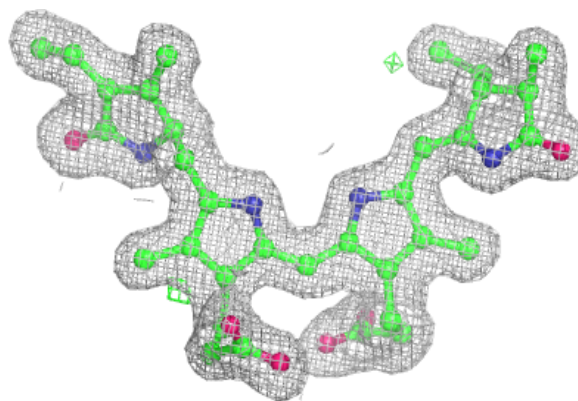
**Electron density around PEB D 203:**

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



**Electron density around PEB B 202:**

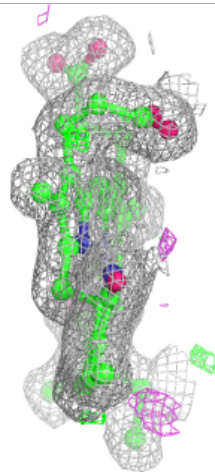
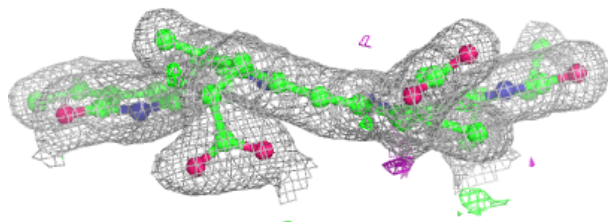
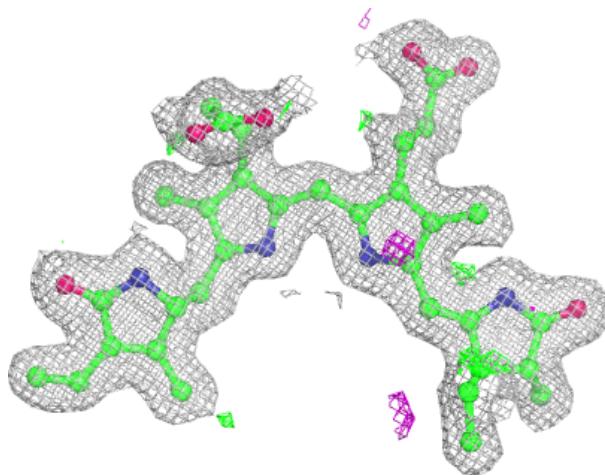
$2mF_o-DF_c$  (at 0.7 rnsd) in gray  
 $mF_o-DF_c$  (at 3 rnsd) in purple (negative)  
and green (positive)





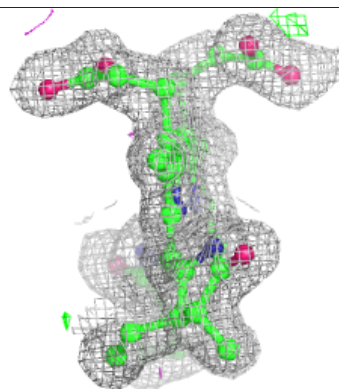
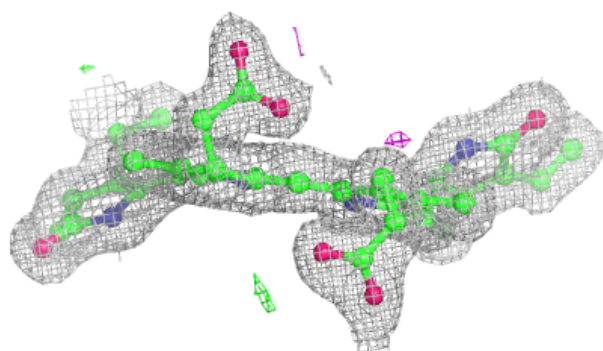
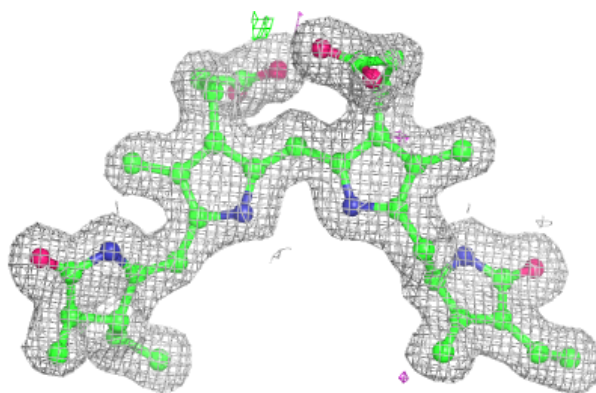
**Electron density around PEB B 203:**

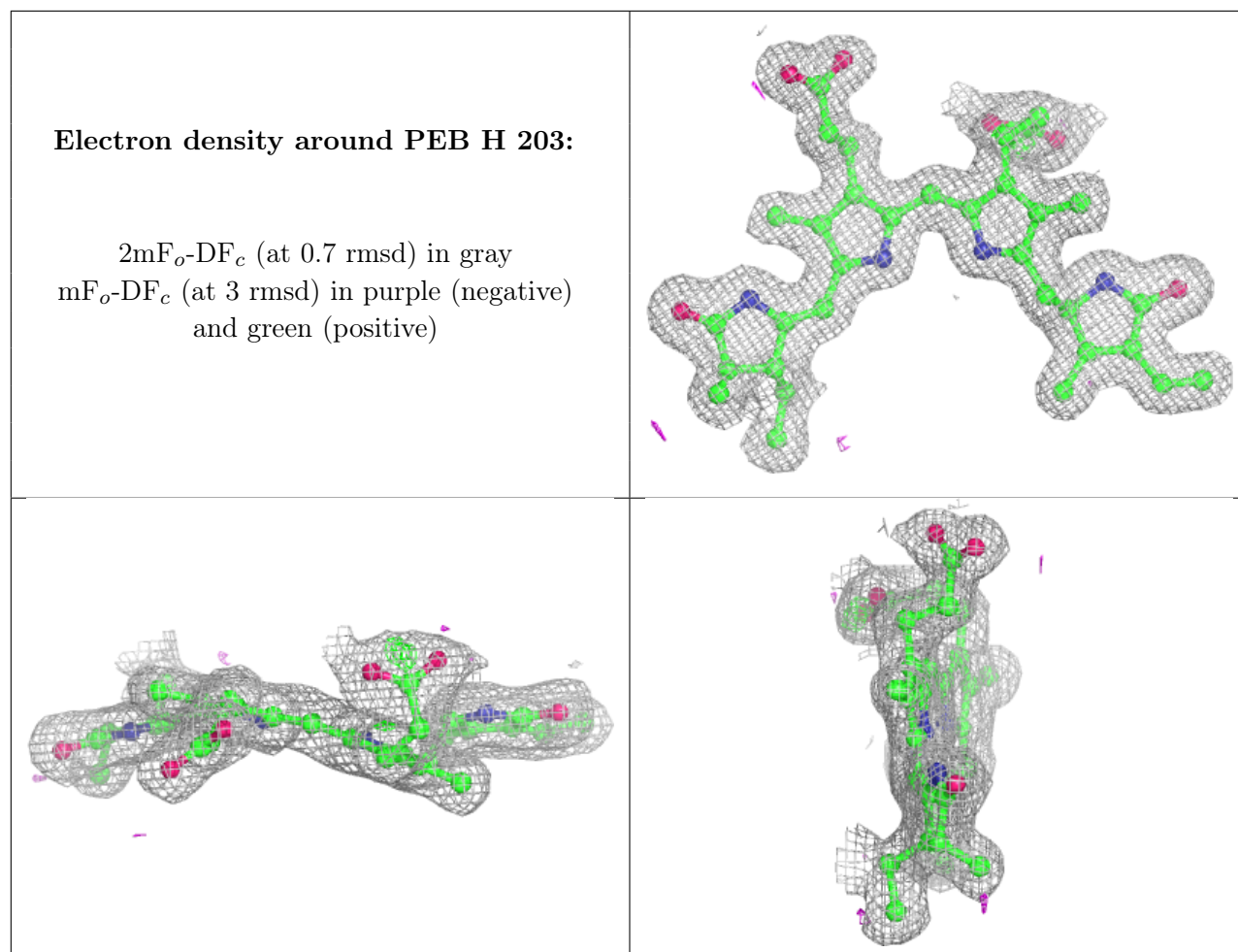
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around PEB F 202:**

$2mF_o-DF_c$  (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.