



## wwPDB EM Validation Summary Report ⓘ

Dec 13, 2022 – 04:59 PM JST

PDB ID : 7DI8  
Title : Electron crystallographic structure of Catalase using a direct electron detector at 300 kV  
Authors : Takaba, K.; Maki-Yonekura, S.; Yonekura, K.  
Deposited on : 2020-11-18  
Resolution : 3.20 Å (reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB/EMDB 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/EMValidationReportHelp>  
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)  
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) : 2.31.3

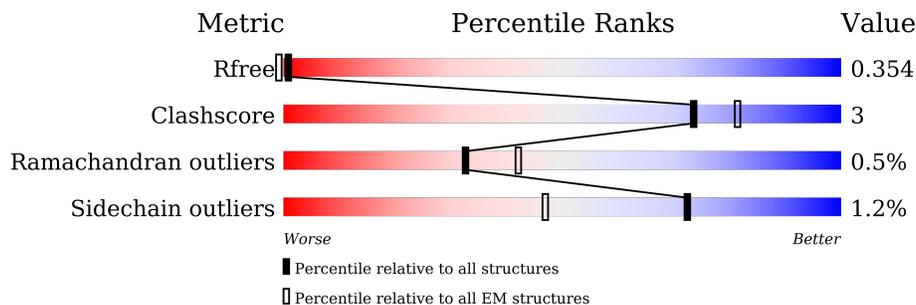
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON CRYSTALLOGRAPHY*

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)	EM structures (#Entries)
$R_{free}$	130704	0
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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  $\leq 5\%$ .

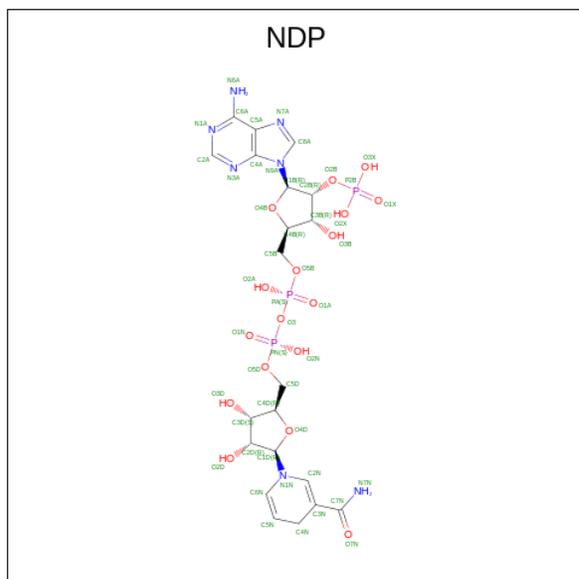
Mol	Chain	Length	Quality of chain
1	A	527	86% 8% 5%
1	B	527	83% 11% 5%
1	C	527	86% 9% 5%
1	D	527	87% 8% 5%



Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	
			Total	C	Fe	H	N		O
2	C	1	Total	C	Fe	H	N	O	0
			73	34	1	30	4	4	
2	D	1	Total	C	Fe	H	N	O	0
			73	34	1	30	4	4	

- Molecule 3 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: C<sub>21</sub>H<sub>30</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>).



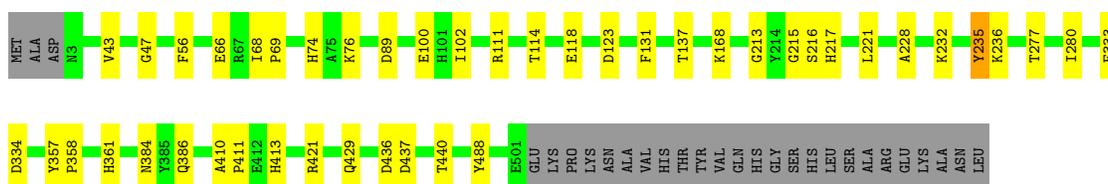
Mol	Chain	Residues	Atoms					AltConf	
			Total	C	H	N	O		P
3	A	1	Total	C	H	N	O	P	0
			74	21	26	7	17	3	
3	B	1	Total	C	H	N	O	P	0
			74	21	26	7	17	3	
3	C	1	Total	C	H	N	O	P	0
			74	21	26	7	17	3	
3	D	1	Total	C	H	N	O	P	0
			74	21	26	7	17	3	

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

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

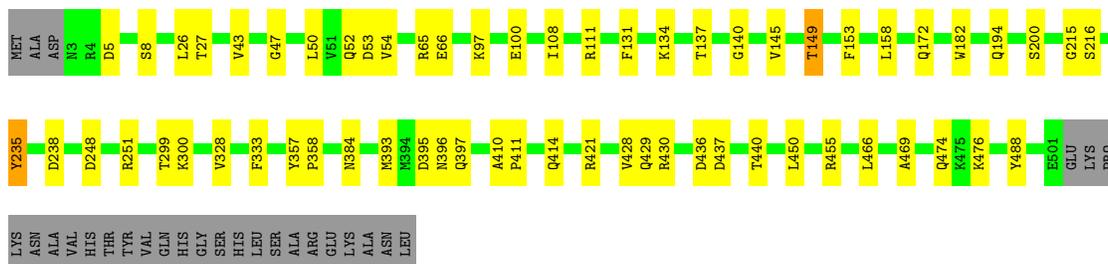
- Molecule 1: Catalase

Chain A: 



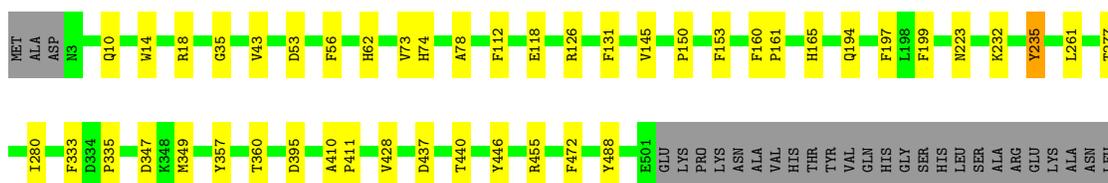
- Molecule 1: Catalase

Chain B: 



- Molecule 1: Catalase

Chain C: 



- Molecule 1: Catalase

Chain D: 



N384	D388	G389	M393	K394	D395	A410	P411	D437	T440	Y488	K498	E501	GLU	GLY	PRO	LYS	LYS	LYS	ASN	ALA	VAL	HIS	THR	THR	TYR	VAL	GLN	HIS	GLY	SER	SER	HIS	LEU	SER	ALA	ARG	GLU	GLY	LYS	ALA	ASN	LEU
------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.98Å 174.02Å 199.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	131.13 – 3.20 131.13 – 3.00	Depositor EDS
% Data completeness (in resolution range)	82.2 (131.13-3.20) 76.0 (131.13-3.00)	Depositor EDS
$R_{merge}$	1.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.79 (at 3.01Å)	Xtrriage
Refinement program	PHENIX 1.12_2829	Depositor
R, $R_{free}$	0.309 , 0.348 0.317 , 0.354	Depositor DCC
$R_{free}$ test set	2000 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	1.9	Xtrriage
Anisotropy	4.568	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.21 , 39.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.32$ , $\langle L^2 \rangle = 0.16$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.51	EDS
Total number of atoms	32035	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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: HEM, NDP

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	0.27	0/4137	0.47	0/5619
1	B	0.27	0/4137	0.46	0/5619
1	C	0.27	0/4137	0.46	0/5619
1	D	0.27	0/4137	0.47	0/5619
All	All	0.27	0/16548	0.47	0/22476

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4017	3846	3843	27	0
1	B	4017	3844	3844	39	0
1	C	4017	3844	3844	32	0
1	D	4017	3845	3843	26	0
2	A	43	30	30	0	0
2	B	43	30	30	0	0
2	C	43	30	30	2	0
2	D	43	30	30	1	0
3	A	48	26	26	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	48	26	26	0	0
3	C	48	26	26	2	0
3	D	48	26	26	1	0
All	All	16432	15603	15598	105	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 105 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:111:ARG:NH1	1:B:328:VAL:O	2.18	0.76
1:B:66:GLU:OE2	1:C:165:HIS:NE2	2.17	0.75
1:B:437:ASP:OD2	1:B:440:THR:OG1	2.08	0.68
1:A:413:HIS:ND1	1:C:35:GLY:O	2.25	0.68
1:A:236:LYS:NZ	3:A:602:NDP:O3X	2.25	0.67

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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	497/527 (94%)	446 (90%)	49 (10%)	2 (0%)	34 69
1	B	497/527 (94%)	462 (93%)	32 (6%)	3 (1%)	25 64
1	C	497/527 (94%)	456 (92%)	39 (8%)	2 (0%)	34 69
1	D	497/527 (94%)	459 (92%)	36 (7%)	2 (0%)	34 69
All	All	1988/2108 (94%)	1823 (92%)	156 (8%)	9 (0%)	32 67

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	384	ASN
1	B	53	ASP
1	C	118	GLU
1	A	216	SER
1	B	395	ASP

### 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 PDB entries followed by that with respect to all EM entries.

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	431/454 (95%)	425 (99%)	6 (1%)	67 86
1	B	431/454 (95%)	426 (99%)	5 (1%)	71 88
1	C	431/454 (95%)	427 (99%)	4 (1%)	78 91
1	D	431/454 (95%)	425 (99%)	6 (1%)	67 86
All	All	1724/1816 (95%)	1703 (99%)	21 (1%)	72 88

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	488	TYR
1	D	125	VAL
1	D	488	TYR
1	D	131	PHE
1	D	104	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	254	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](#)

8 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
3	NDP	A	602	-	45,52,52	2.28	6 (13%)	53,80,80	1.62	9 (16%)
2	HEM	B	601	1	41,50,50	1.48	6 (14%)	45,82,82	1.49	7 (15%)
3	NDP	B	602	-	45,52,52	2.43	6 (13%)	53,80,80	1.66	9 (16%)
3	NDP	D	602	-	45,52,52	2.36	6 (13%)	53,80,80	1.65	10 (18%)
2	HEM	D	601	1	41,50,50	1.51	6 (14%)	45,82,82	1.52	8 (17%)
3	NDP	C	602	-	45,52,52	2.35	6 (13%)	53,80,80	1.68	8 (15%)
2	HEM	A	601	1	41,50,50	1.52	5 (12%)	45,82,82	1.48	7 (15%)
2	HEM	C	601	1	41,50,50	1.42	3 (7%)	45,82,82	1.55	6 (13%)

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
3	NDP	A	602	-	-	9/30/77/77	0/5/5/5
2	HEM	B	601	1	-	4/12/54/54	-
3	NDP	B	602	-	-	10/30/77/77	0/5/5/5
3	NDP	D	602	-	-	7/30/77/77	0/5/5/5

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEM	D	601	1	-	2/12/54/54	-
3	NDP	C	602	-	-	5/30/77/77	0/5/5/5
2	HEM	A	601	1	-	2/12/54/54	-
2	HEM	C	601	1	-	4/12/54/54	-

The worst 5 of 44 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	602	NDP	P2B-O2B	13.46	1.84	1.59
3	D	602	NDP	P2B-O2B	13.15	1.84	1.59
3	C	602	NDP	P2B-O2B	12.87	1.83	1.59
3	A	602	NDP	P2B-O2B	12.31	1.82	1.59
3	C	602	NDP	PN-O5D	4.70	1.78	1.59

The worst 5 of 64 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	602	NDP	PN-O3-PA	-7.29	107.82	132.83
3	C	602	NDP	PN-O3-PA	-7.14	108.33	132.83
3	A	602	NDP	PN-O3-PA	-6.89	109.19	132.83
3	D	602	NDP	PN-O3-PA	-6.75	109.65	132.83
2	B	601	HEM	C4C-CHD-C1D	3.63	127.35	122.56

There are no chirality outliers.

5 of 43 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	602	NDP	C5B-O5B-PA-O1A
3	B	602	NDP	C5B-O5B-PA-O1A
3	B	602	NDP	C5B-O5B-PA-O2A
3	B	602	NDP	C2N-C3N-C7N-N7N
3	D	602	NDP	C2N-C3N-C7N-N7N

There are no ring outliers.

5 monomers are involved in 7 short contacts:

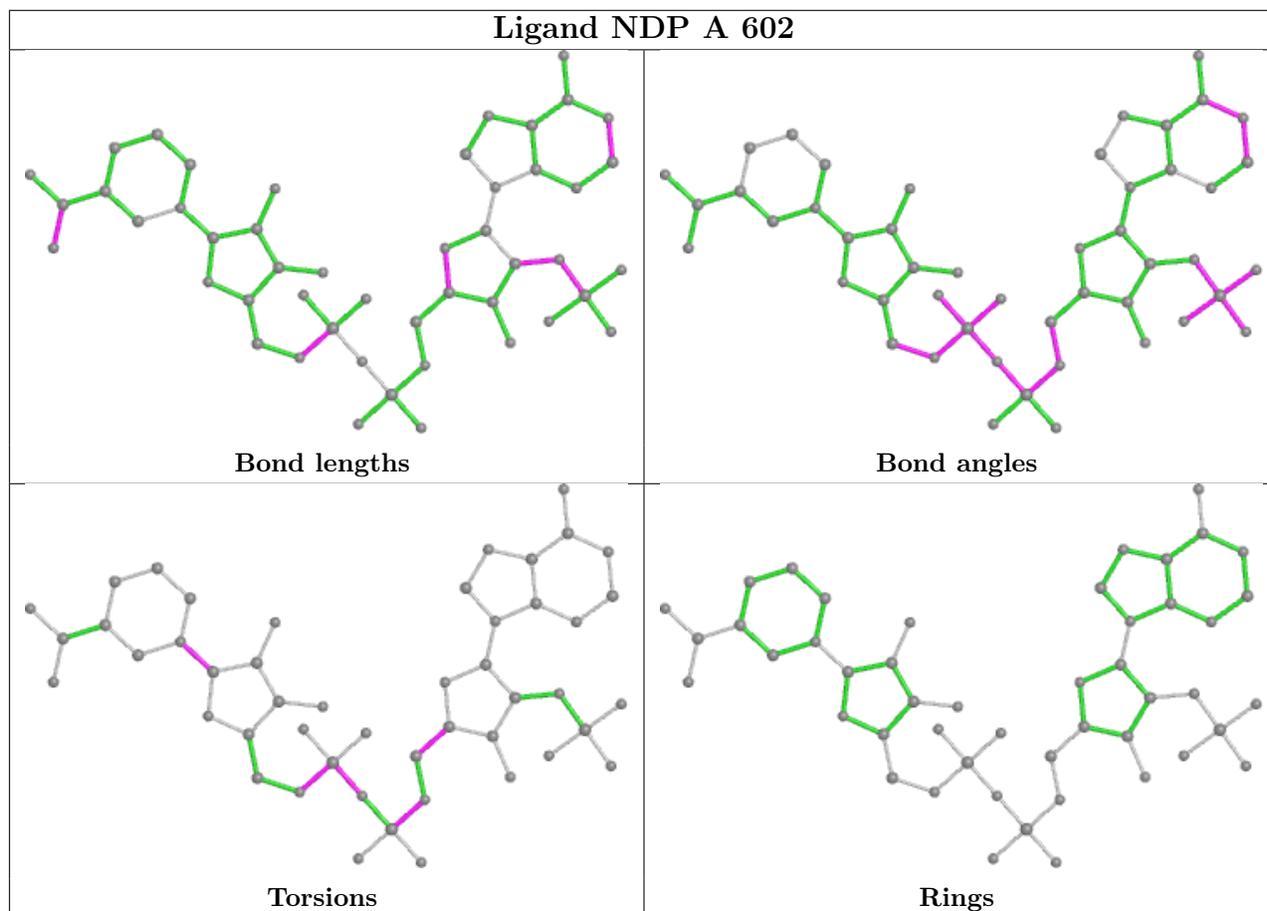
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	602	NDP	1	0
3	D	602	NDP	1	0
2	D	601	HEM	1	0

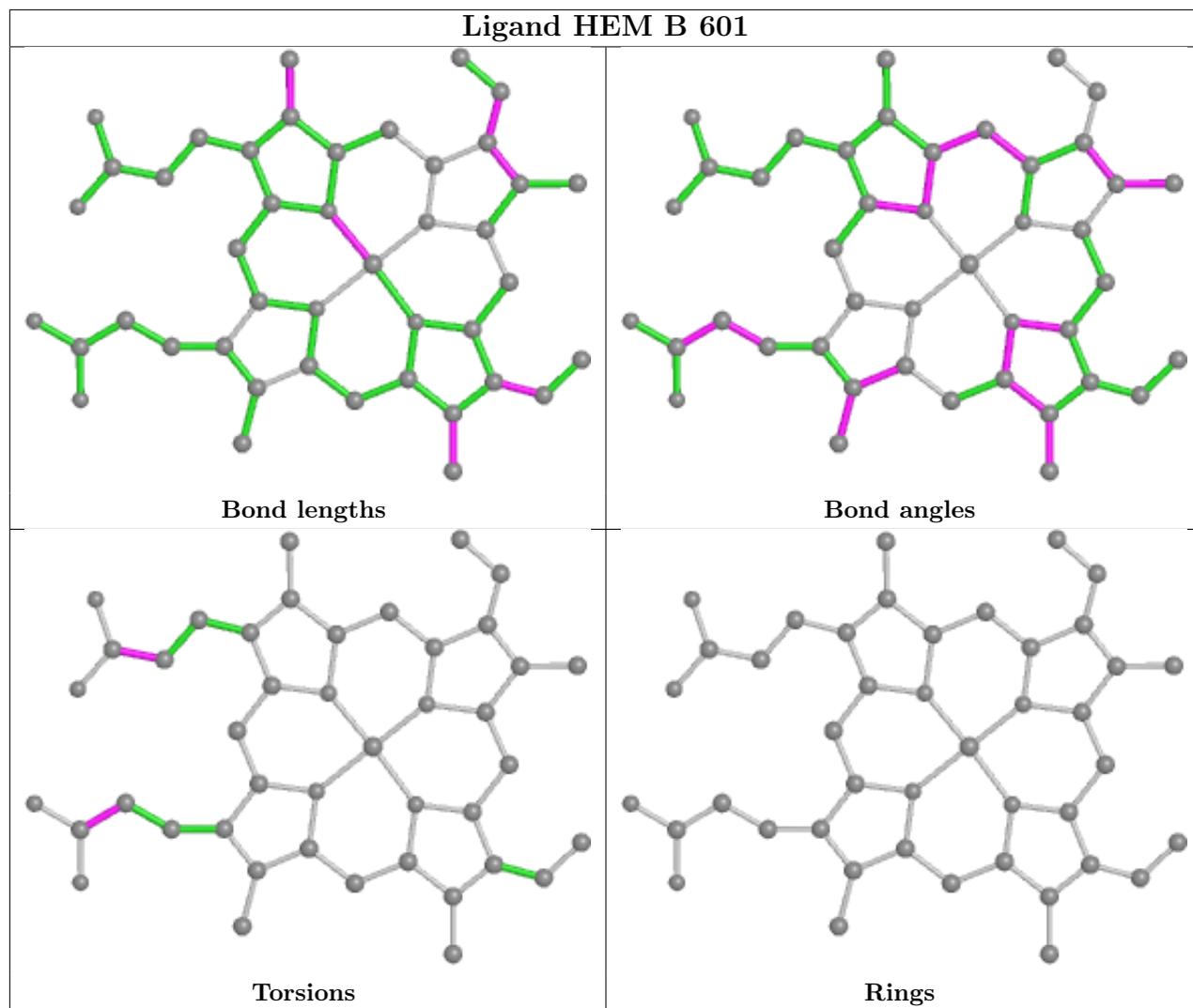
*Continued on next page...*

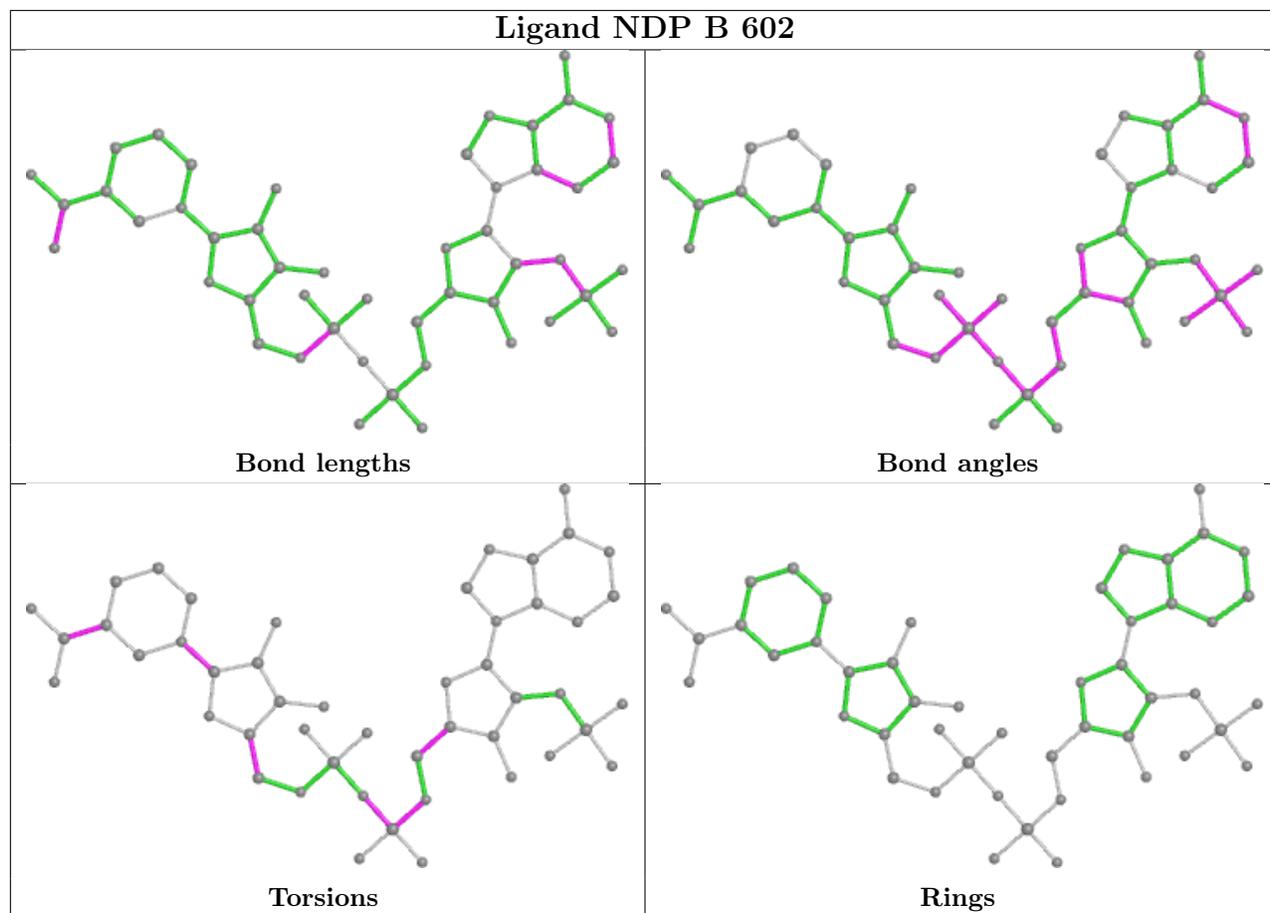
Continued from previous page...

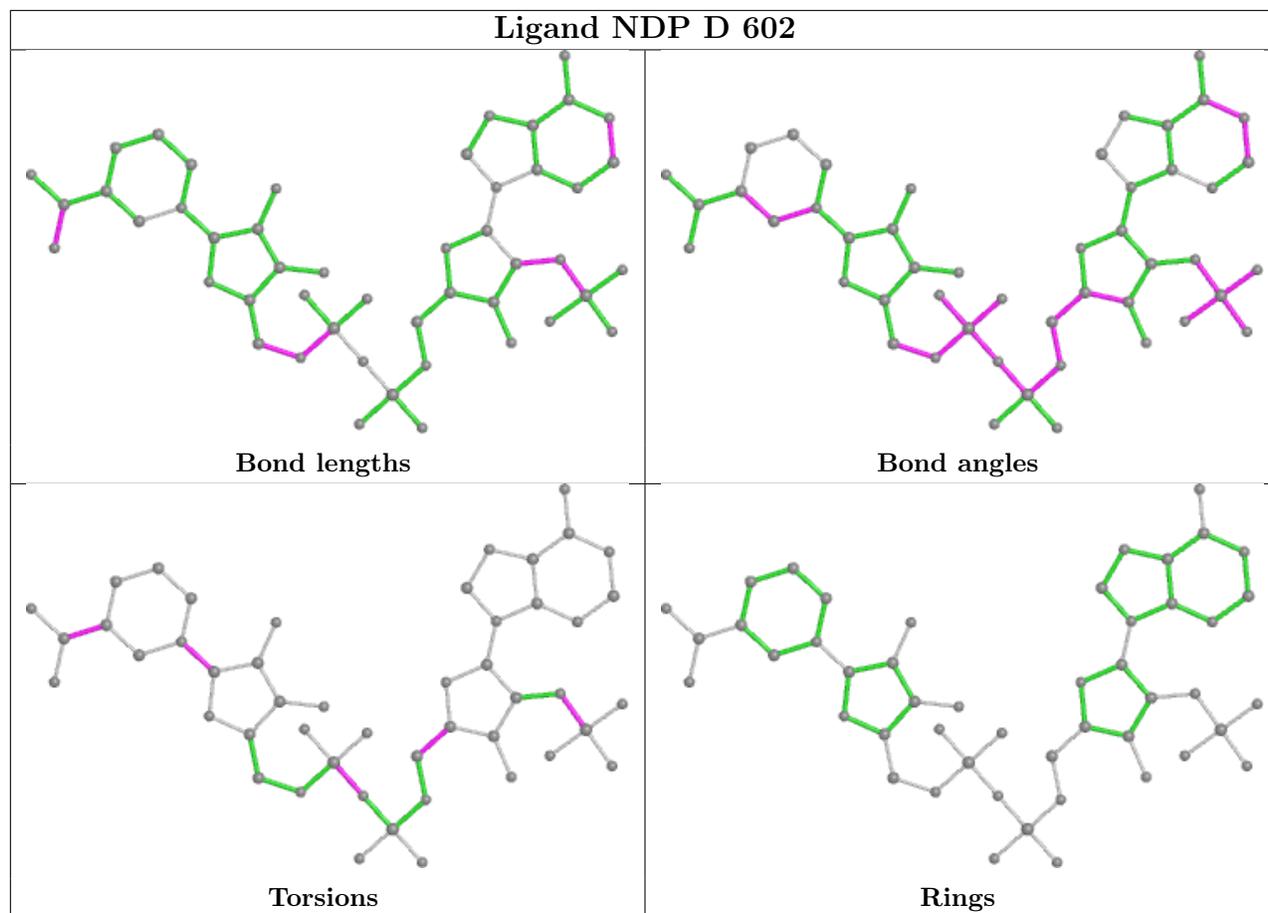
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	602	NDP	2	0
2	C	601	HEM	2	0

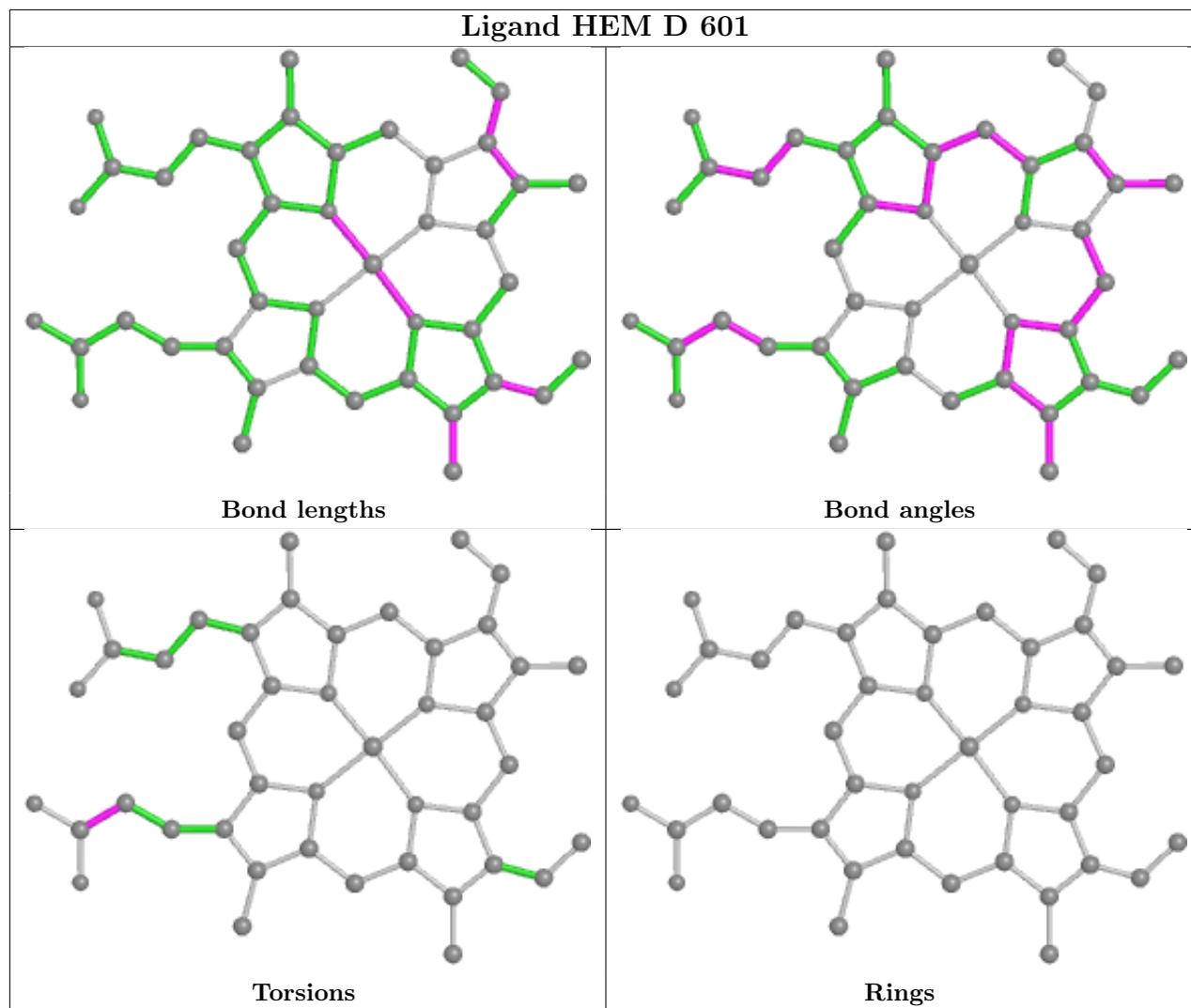
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

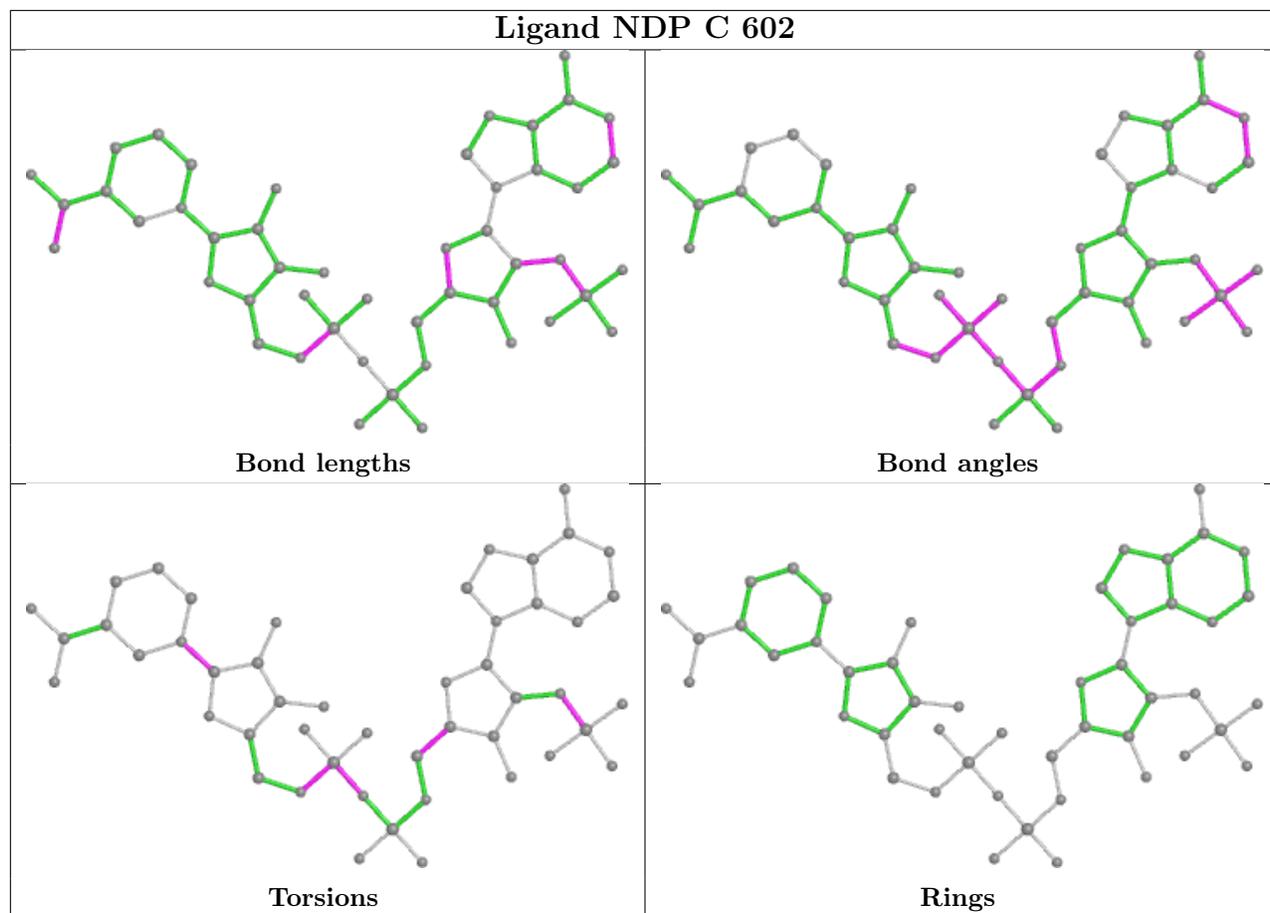


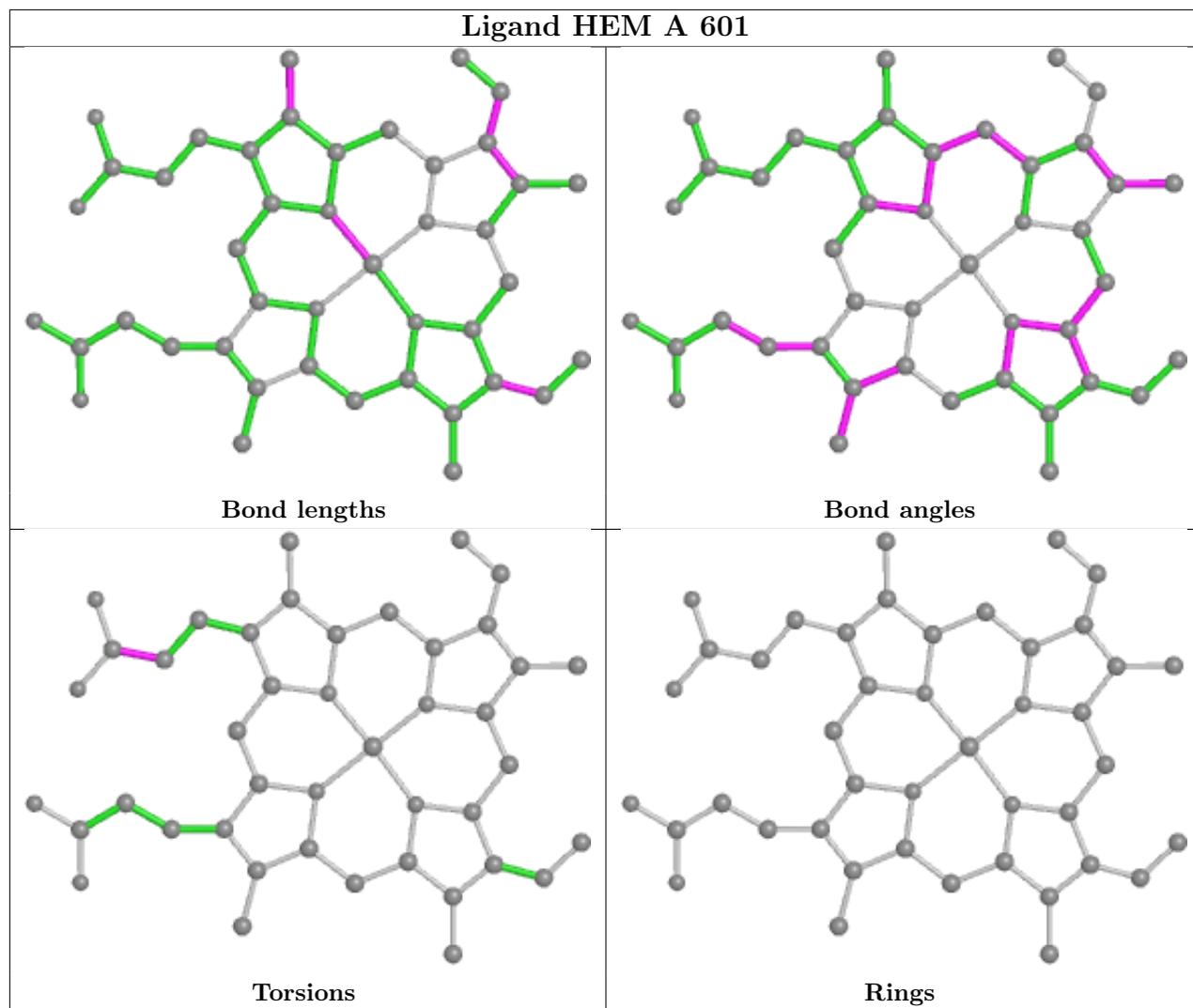


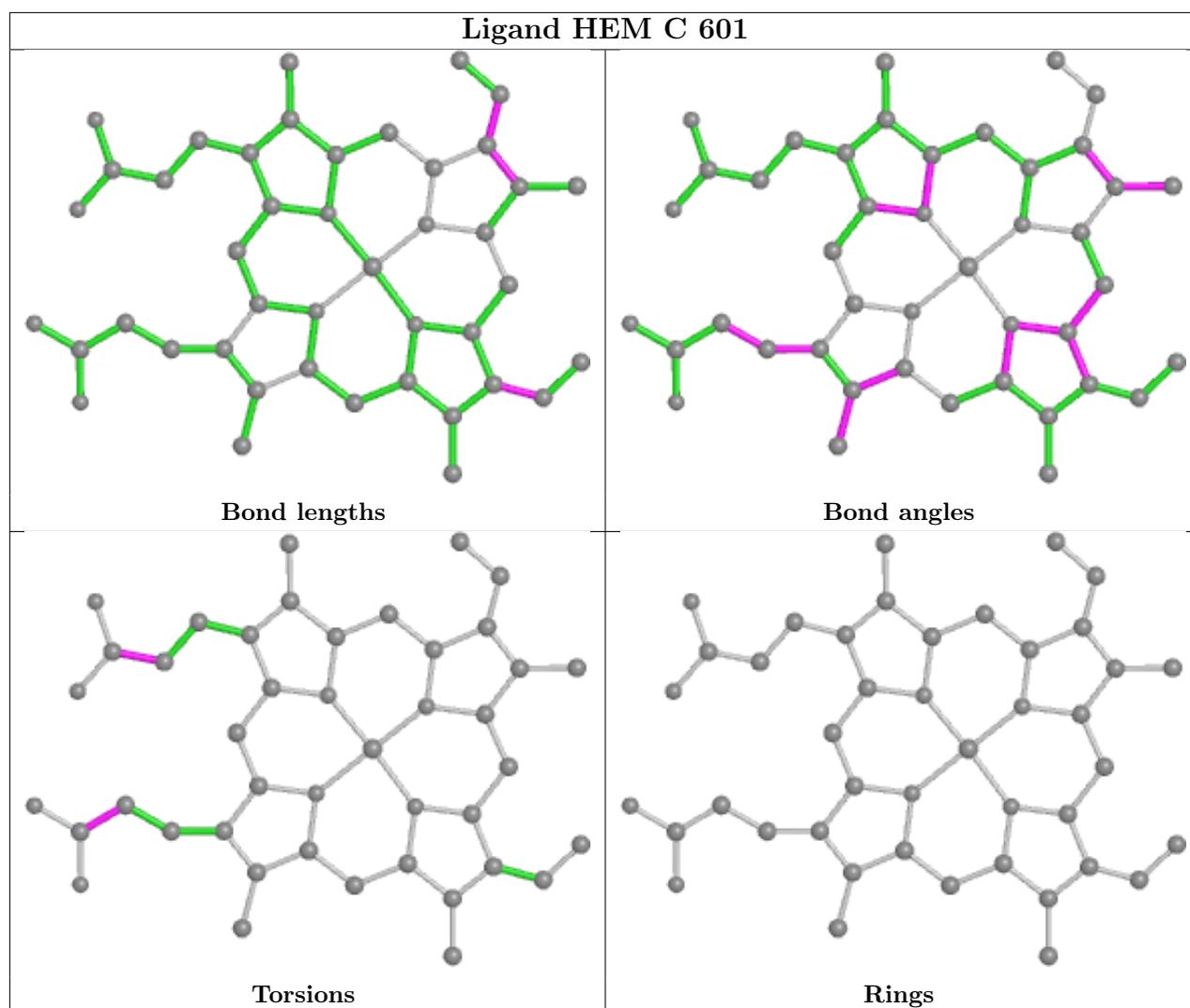












## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.