



## wwPDB EM Validation Summary Report ⓘ

Oct 31, 2023 – 04:47 PM JST

PDB ID : 5GKN  
Title : Catalase structure determined by electron crystallography of thin 3D crystals  
Authors : Yonekura, K.  
Deposited on : 2016-07-04  
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) : 2.36

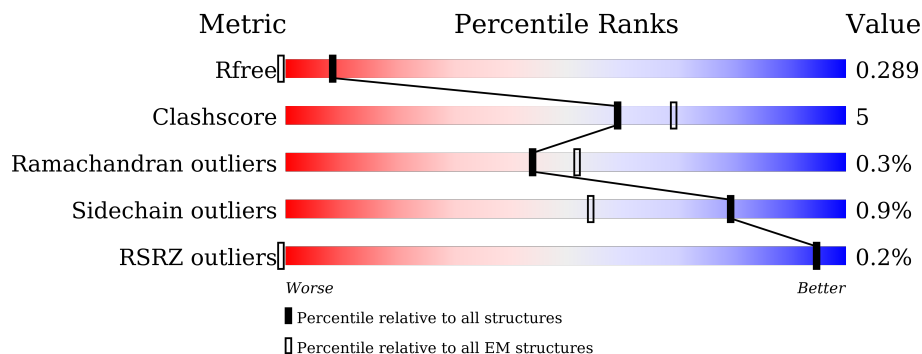
# 1 Overall quality at a glance i

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
RSRZ outliers	127900	0

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	82% 12% • 5%
1	B	527	82% 12% • 5%
1	C	527	85% 9% 5%
1	D	527	84% 10% • 5%

## 2 Entry composition [i](#)

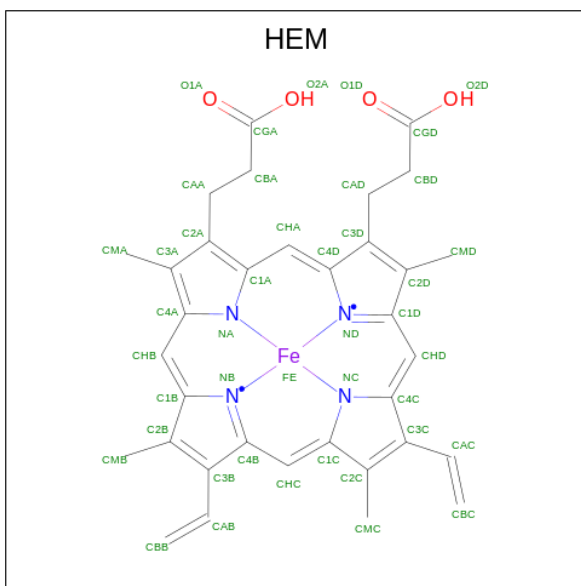
There are 4 unique types of molecules in this entry. The entry contains 16548 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Catalase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	499	Total	C	N	O	S	0	0
			4017	2548	717	738	14		
1	B	499	Total	C	N	O	S	0	0
			4017	2548	717	738	14		
1	C	499	Total	C	N	O	S	0	0
			4017	2548	717	738	14		
1	D	499	Total	C	N	O	S	0	0
			4017	2548	717	738	14		

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



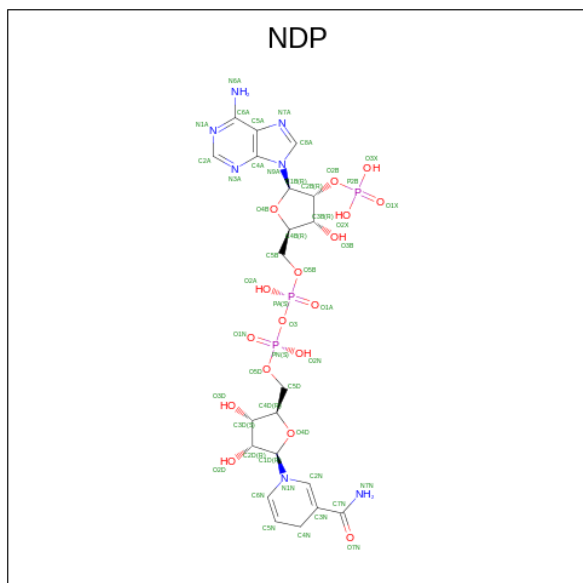
Mol	Chain	Residues	Atoms					AltConf
			Total	C	Fe	N	O	
2	A	1	Total	C	Fe	N	O	0
			43	34	1	4	4	
2	B	1	Total	C	Fe	N	O	0
			43	34	1	4	4	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				AltConf	
2	C	1	Total	C	Fe	N	O	0
			43	34	1	4	4	
2	D	1	Total	C	Fe	N	O	0
			43	34	1	4	4	

- Molecule 3 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula:  $C_{21}H_{30}N_7O_{17}P_3$ ).



Mol	Chain	Residues	Atoms				AltConf	
3	A	1	Total	C	N	O	P	0
			48	21	7	17	3	
3	B	1	Total	C	N	O	P	0
			48	21	7	17	3	
3	C	1	Total	C	N	O	P	0
			48	21	7	17	3	
3	D	1	Total	C	N	O	P	0
			48	21	7	17	3	

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		AltConf
4	A	35	Total	O	0
			35	35	
4	B	18	Total	O	0
			18	18	
4	C	30	Total	O	0
			30	30	

Continued on next page...


*Continued from previous page...*

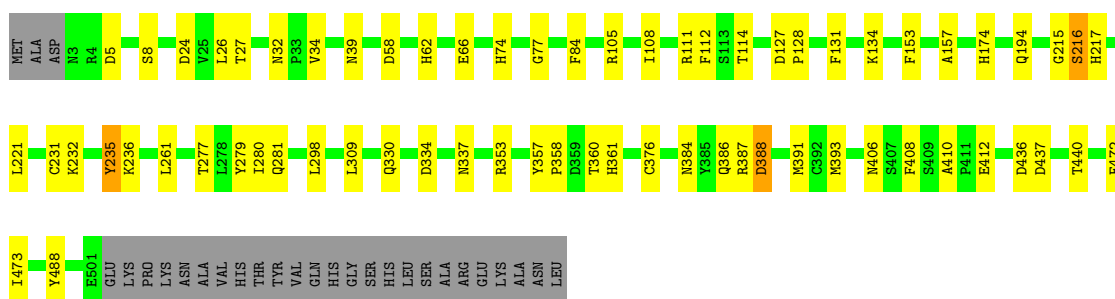
Mol	Chain	Residues	Atoms		AltConf
4	D	33	Total	O	0
			33	33	

### 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.

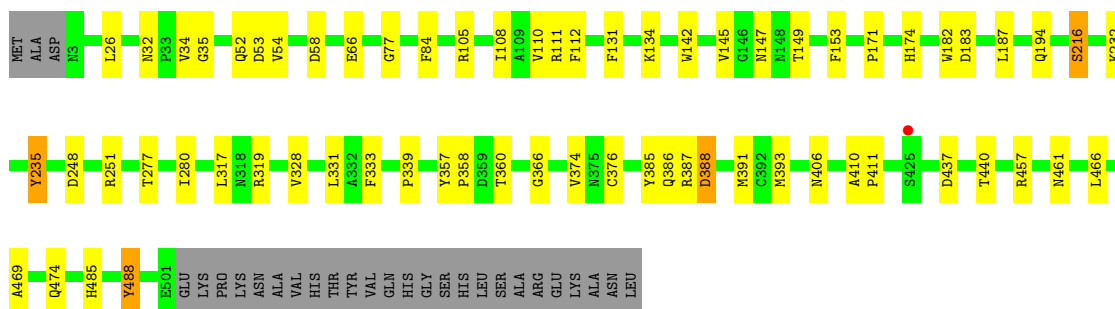
- Molecule 1: Catalase

Chain A: 




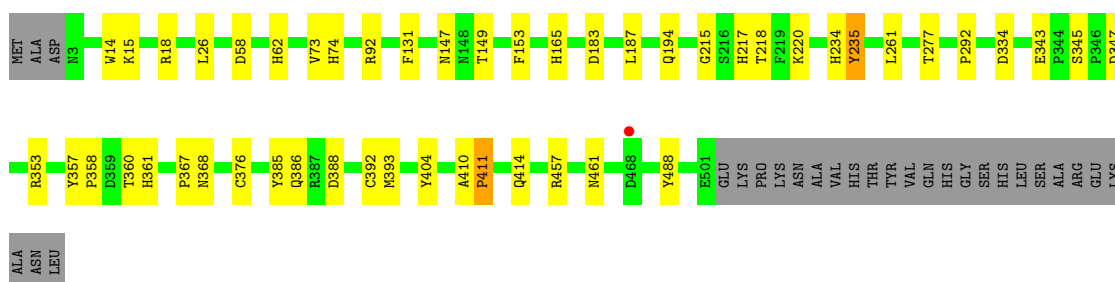
- Molecule 1: Catalase

Chain B: 




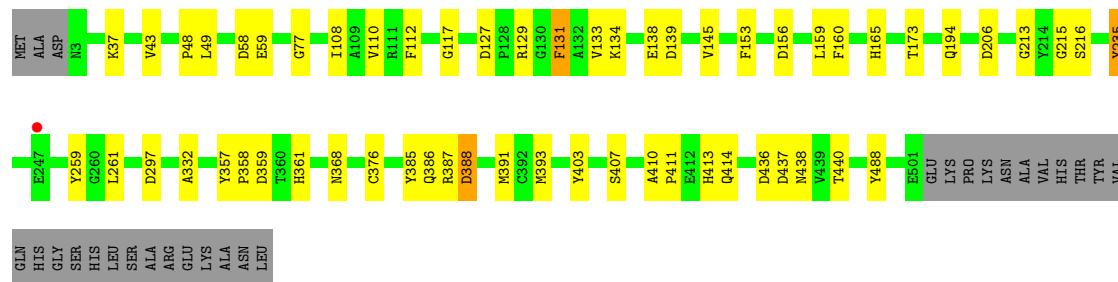
- Molecule 1: Catalase

Chain C: 



- Molecule 1: Catalase

Chain D:  84% 10% • 5%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.00Å 173.50Å 206.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 3.20 19.99 – 3.20	Depositor EDS
% Data completeness (in resolution range)	73.0 (19.99-3.20) 73.0 (19.99-3.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	17.46 (at 3.22Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.251 , 0.304 0.267 , 0.289	Depositor DCC
$R_{free}$ test set	893 reflections (2.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.0	Xtrriage
Anisotropy	1.344	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.15 , 41.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.21$ , $\langle L^2 \rangle = 0.07$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.65	EDS
Total number of atoms	16548	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 3.27% 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	0.38	1/4137 (0.0%)	0.57	0/5619
1	B	0.41	1/4137 (0.0%)	0.54	0/5619
1	C	0.38	1/4137 (0.0%)	0.54	0/5619
1	D	0.38	1/4137 (0.0%)	0.54	0/5619
All	All	0.39	4/16548 (0.0%)	0.55	0/22476

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	376	CYS	C-N	13.35	1.59	1.34
1	D	376	CYS	C-N	8.37	1.50	1.34
1	C	376	CYS	C-N	-7.34	1.20	1.34
1	A	376	CYS	C-N	6.44	1.46	1.34

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	0	3843	49	0
1	B	4017	0	3843	51	0
1	C	4017	0	3843	41	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	4017	0	3843	47	0
2	A	43	0	30	4	0
2	B	43	0	30	4	0
2	C	43	0	30	5	0
2	D	43	0	30	1	0
3	A	48	0	26	1	0
3	B	48	0	26	0	0
3	C	48	0	26	0	0
3	D	48	0	26	0	0
4	A	35	0	0	1	0
4	B	18	0	0	3	0
4	C	30	0	0	3	0
4	D	33	0	0	5	0
All	All	16548	0	15596	155	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 5.

The worst 5 of 155 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:C:92:ARG:NH2	4:C:701:HOH:O	1.91	1.03
1:C:367:PRO:O	4:C:702:HOH:O	1.91	0.89
1:B:111:ARG:NH1	1:B:328:VAL:O	2.14	0.81
1:A:386:GLN:NE2	1:C:58:ASP:OD2	2.13	0.81
1:A:66:GLU:OE2	1:D:165:HIS:NE2	2.15	0.80

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%)	465 (94%)	30 (6%)	2 (0%)	34	69
1	B	497/527 (94%)	467 (94%)	28 (6%)	2 (0%)	34	69
1	C	497/527 (94%)	471 (95%)	25 (5%)	1 (0%)	47	79
1	D	497/527 (94%)	465 (94%)	31 (6%)	1 (0%)	47	79
All	All	1988/2108 (94%)	1868 (94%)	114 (6%)	6 (0%)	44	74

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	388	ASP
1	A	216	SER
1	A	388	ASP
1	B	53	ASP
1	D	388	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%)	427 (99%)	4 (1%)	78	91
1	B	431/454 (95%)	427 (99%)	4 (1%)	78	91
1	C	431/454 (95%)	427 (99%)	4 (1%)	78	91
1	D	431/454 (95%)	427 (99%)	4 (1%)	78	91
All	All	1724/1816 (95%)	1708 (99%)	16 (1%)	79	91

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

Mol	Chain	Res	Type
1	D	235	TYR
1	D	159	LEU
1	C	131	PHE
1	D	131	PHE
1	B	488	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	330	GLN
1	B	397	GLN
1	B	461	ASN
1	D	101	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	C	602	-	45,52,52	1.01	3 (6%)	53,80,80	1.49	9 (16%)
2	HEM	B	601	1	41,50,50	1.45	7 (17%)	45,82,82	2.49	19 (42%)
2	HEM	A	601	1	41,50,50	1.42	6 (14%)	45,82,82	1.96	14 (31%)
3	NDP	B	602	-	45,52,52	1.01	3 (6%)	53,80,80	1.53	9 (16%)
3	NDP	D	602	-	45,52,52	0.97	3 (6%)	53,80,80	1.30	8 (15%)
2	HEM	D	601	1	41,50,50	1.38	6 (14%)	45,82,82	2.10	13 (28%)
3	NDP	A	602	-	45,52,52	1.02	2 (4%)	53,80,80	1.29	6 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	HEM	C	601	1	41,50,50	1.43	7 (17%)	45,82,82	2.01	14 (31%)

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	C	602	-	-	5/30/77/77	0/5/5/5
2	HEM	B	601	1	-	4/12/54/54	-
2	HEM	A	601	1	-	4/12/54/54	-
3	NDP	B	602	-	-	12/30/77/77	0/5/5/5
3	NDP	D	602	-	-	8/30/77/77	0/5/5/5
2	HEM	D	601	1	-	2/12/54/54	-
3	NDP	A	602	-	-	6/30/77/77	0/5/5/5
2	HEM	C	601	1	-	4/12/54/54	-

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	HEM	C1B-NB	-4.06	1.33	1.40
2	B	601	HEM	C1B-NB	-4.00	1.33	1.40
2	C	601	HEM	C1B-NB	-3.84	1.33	1.40
2	D	601	HEM	C1B-NB	-3.81	1.33	1.40
3	D	602	NDP	C6N-C5N	3.63	1.39	1.33

The worst 5 of 92 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	601	HEM	CHC-C4B-NB	6.46	131.44	124.43
2	B	601	HEM	CHC-C4B-NB	6.44	131.43	124.43
2	A	601	HEM	CHC-C4B-NB	4.97	129.84	124.43
2	C	601	HEM	CHD-C1D-ND	4.78	129.62	124.43
2	B	601	HEM	CHD-C1D-ND	4.77	129.61	124.43

There are no chirality outliers.

5 of 45 torsion outliers are listed below:

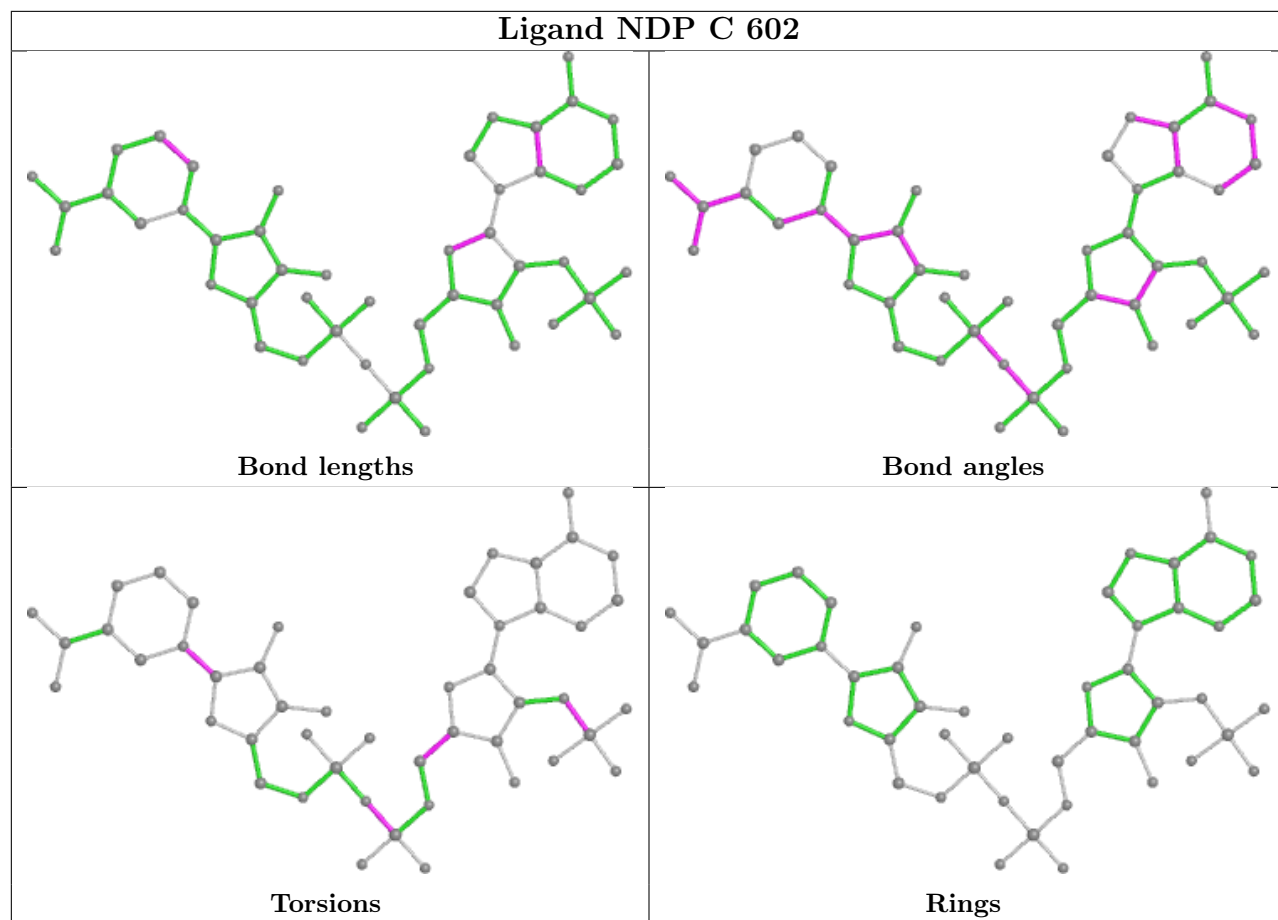
Mol	Chain	Res	Type	Atoms
3	B	602	NDP	C5B-O5B-PA-O1A
3	B	602	NDP	C2B-O2B-P2B-O3X
3	D	602	NDP	C2N-C3N-C7N-N7N
3	B	602	NDP	O4B-C4B-C5B-O5B
3	B	602	NDP	C3B-C4B-C5B-O5B

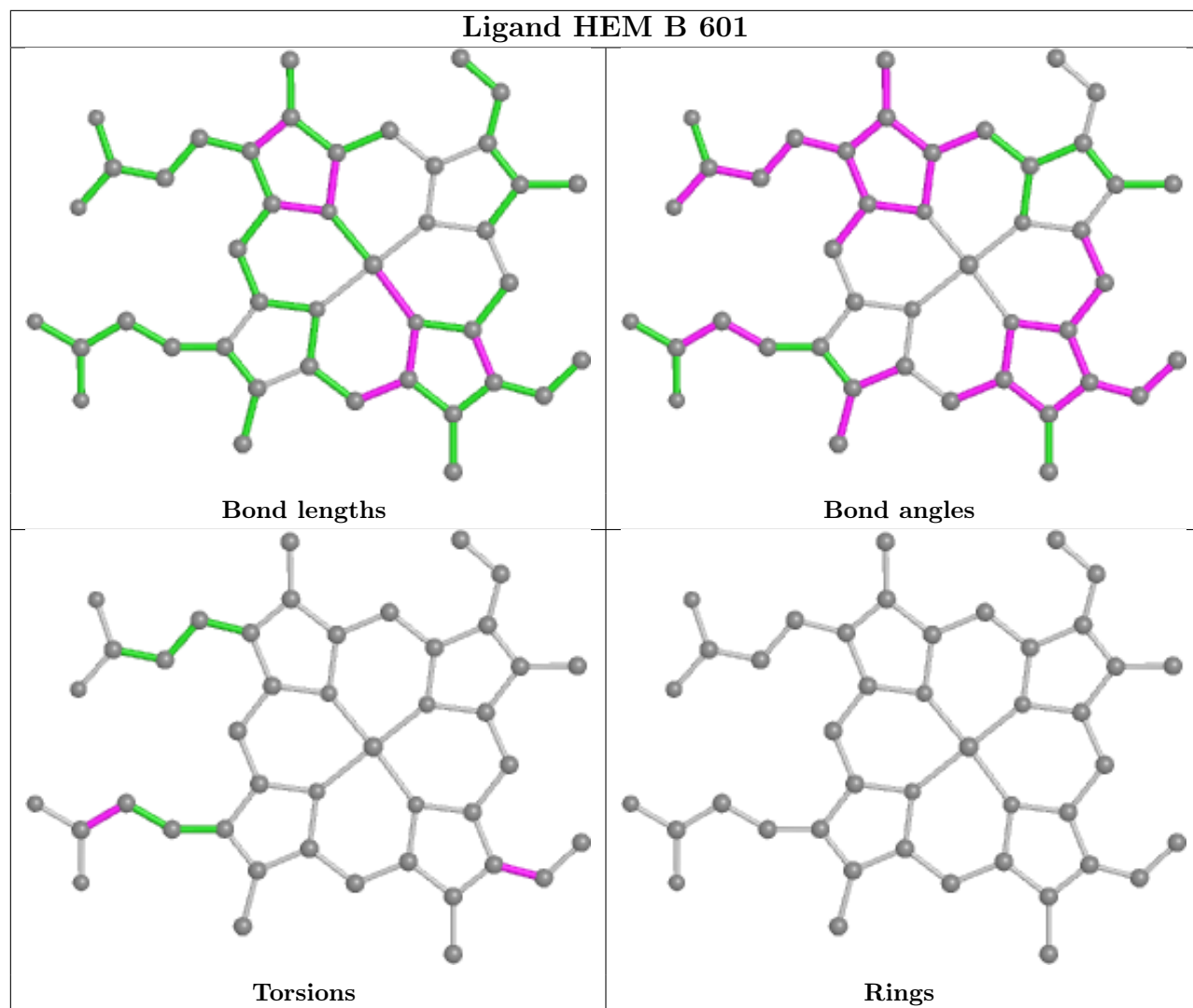
There are no ring outliers.

5 monomers are involved in 15 short contacts:

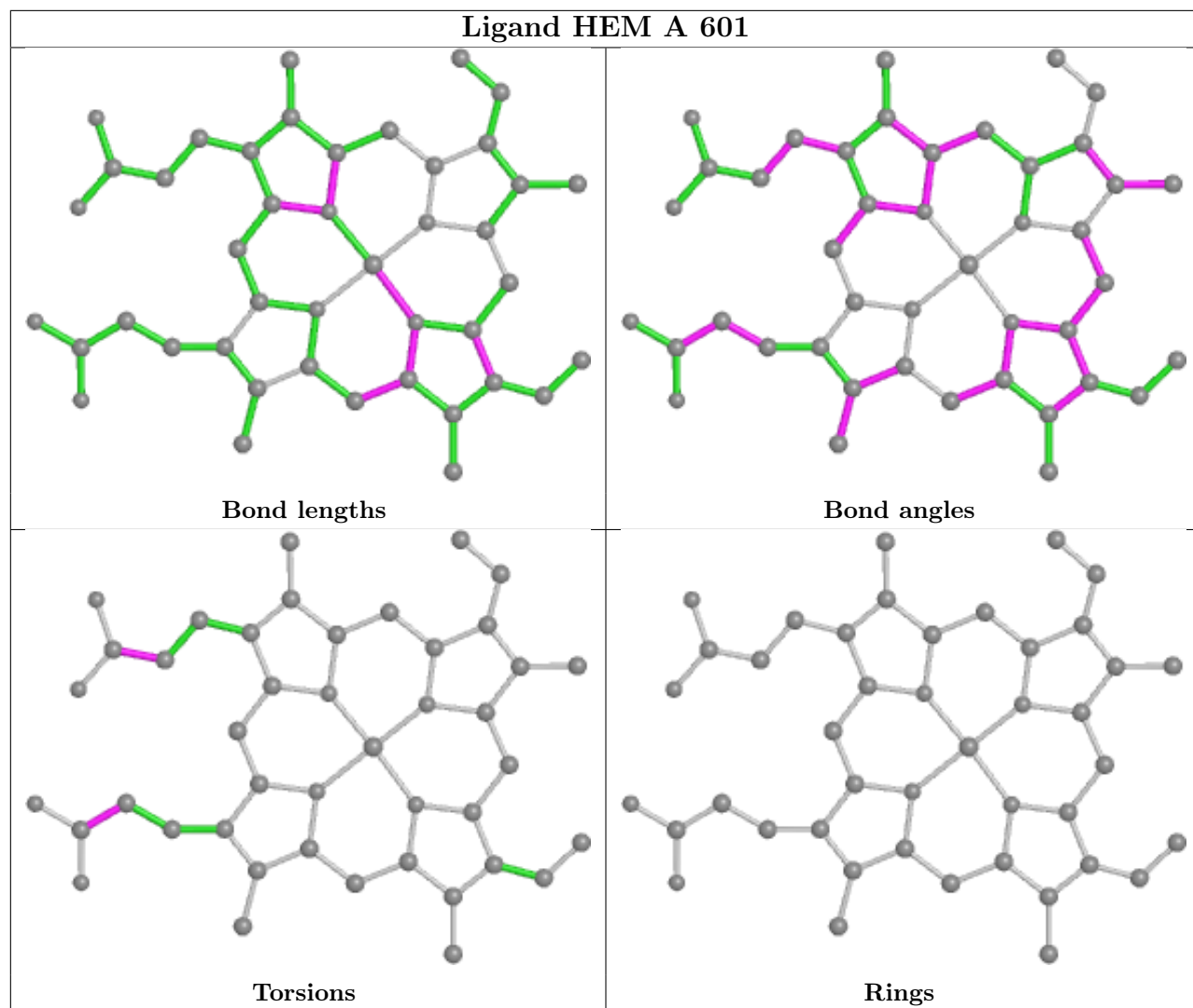
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	601	HEM	4	0
2	A	601	HEM	4	0
2	D	601	HEM	1	0
3	A	602	NDP	1	0
2	C	601	HEM	5	0

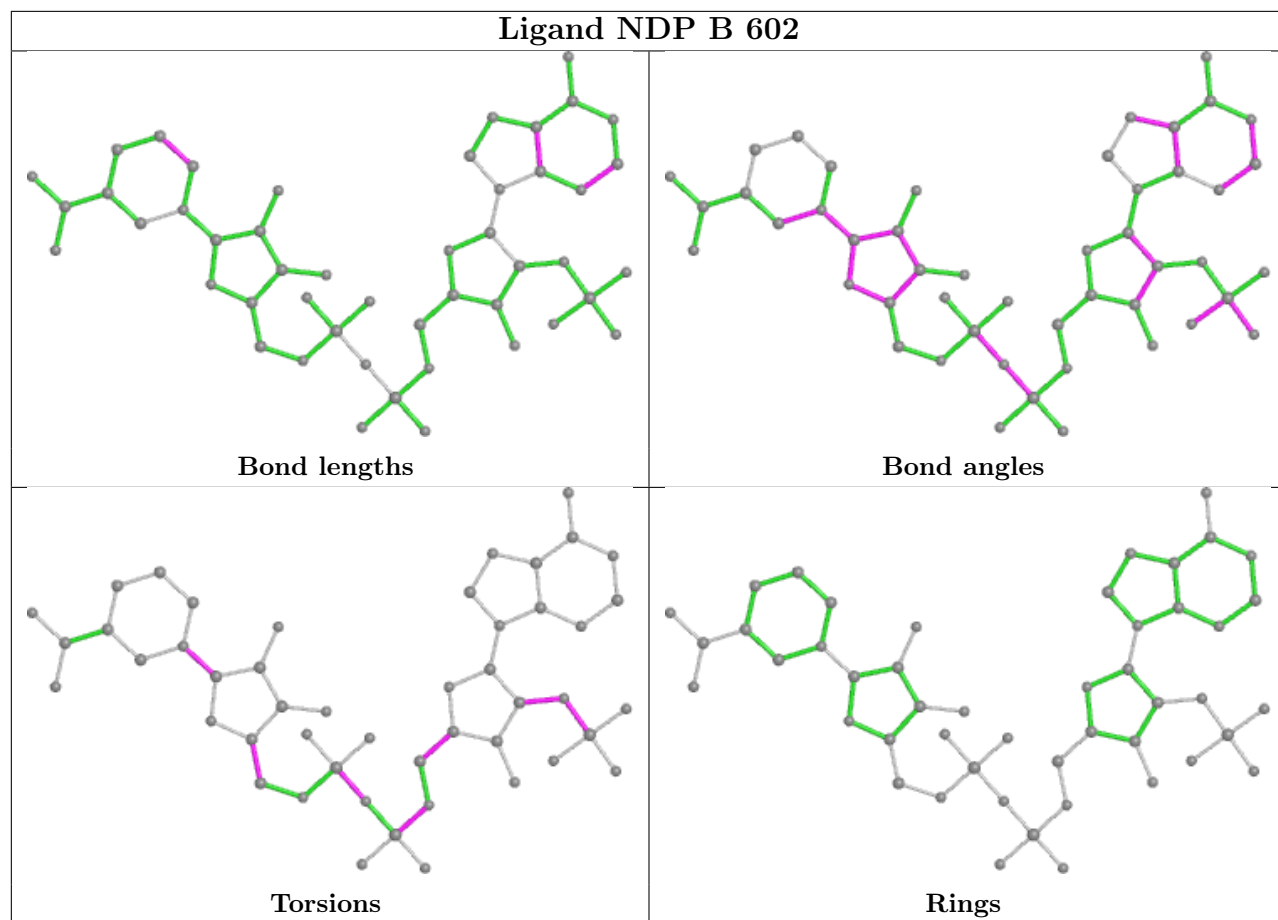
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

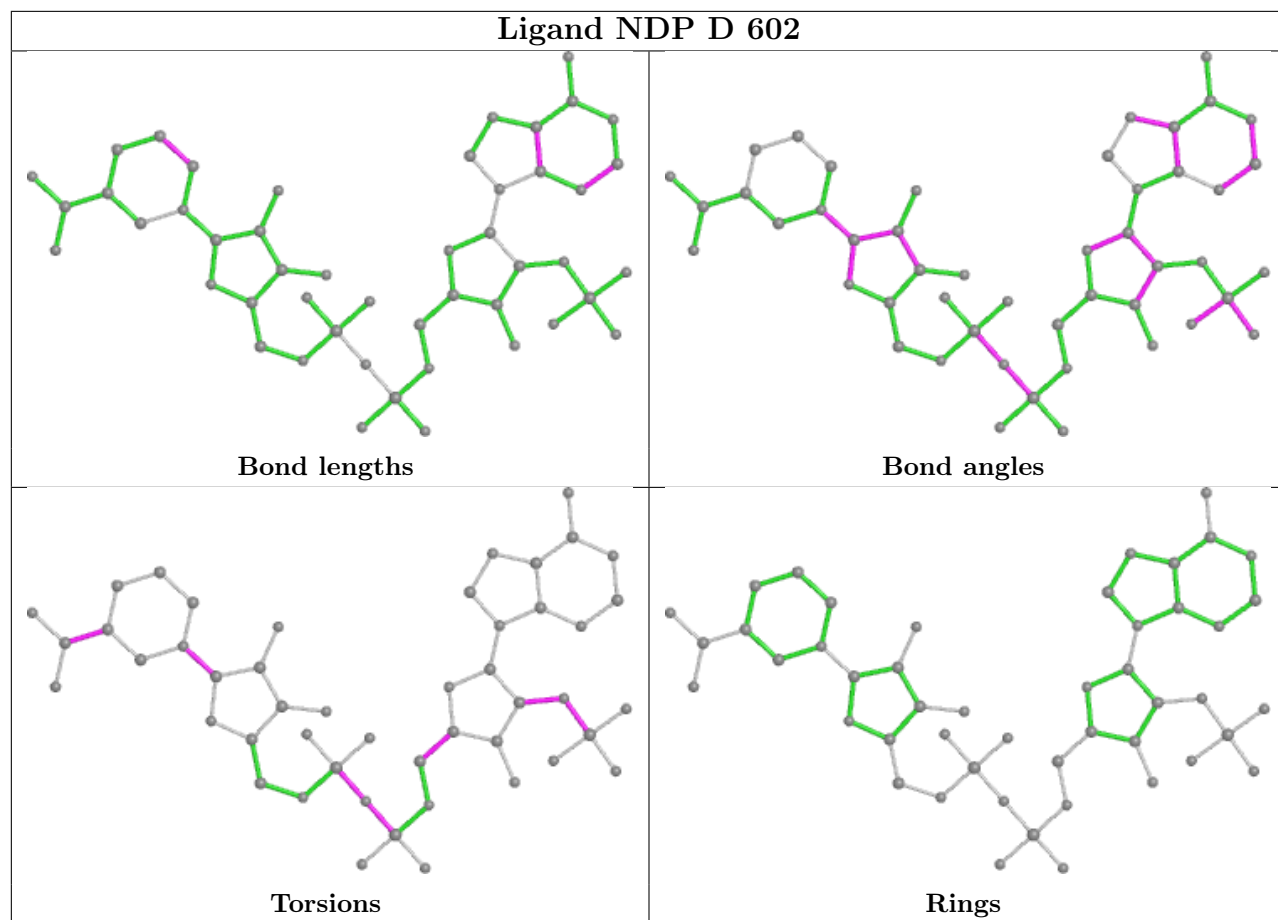


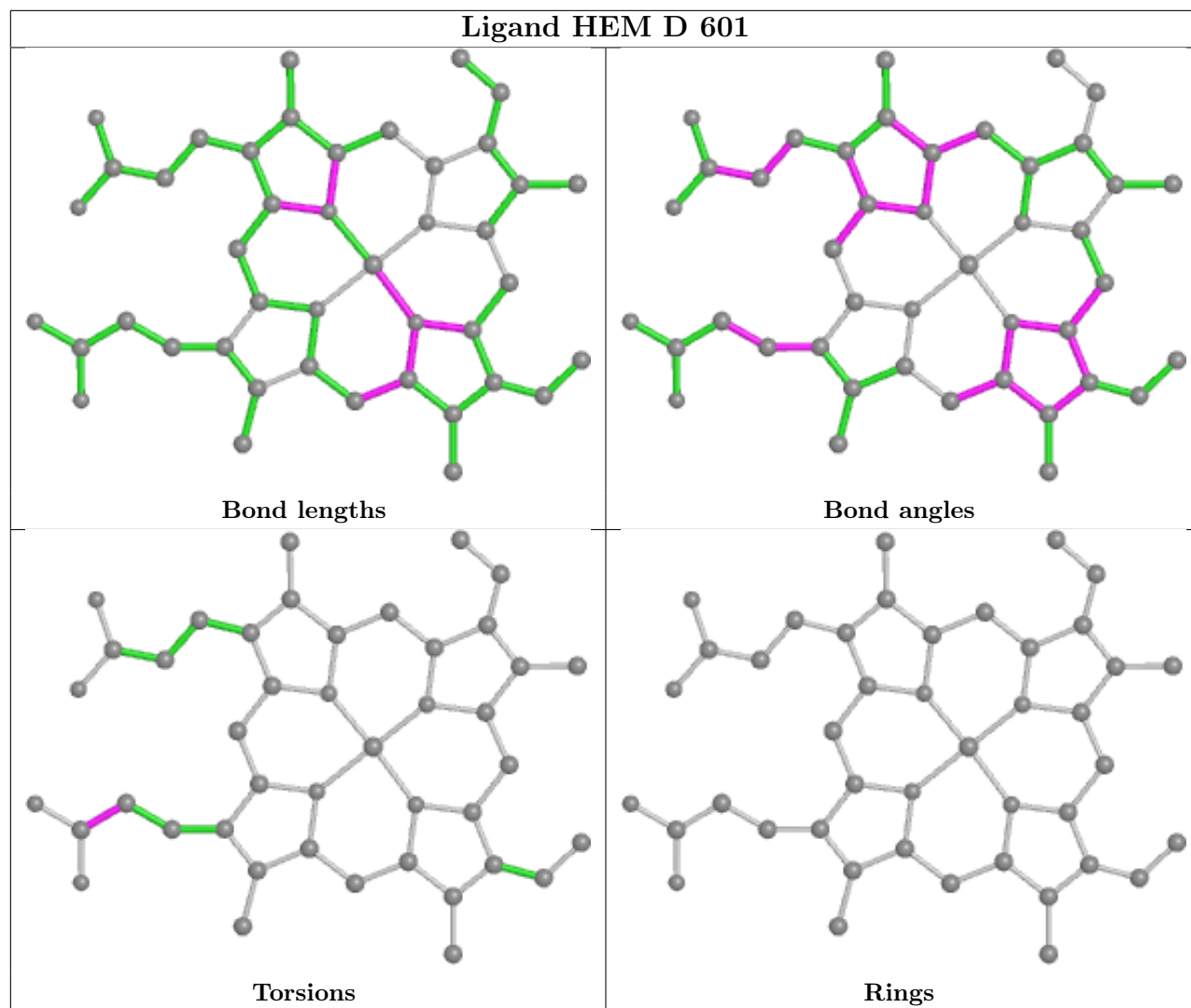


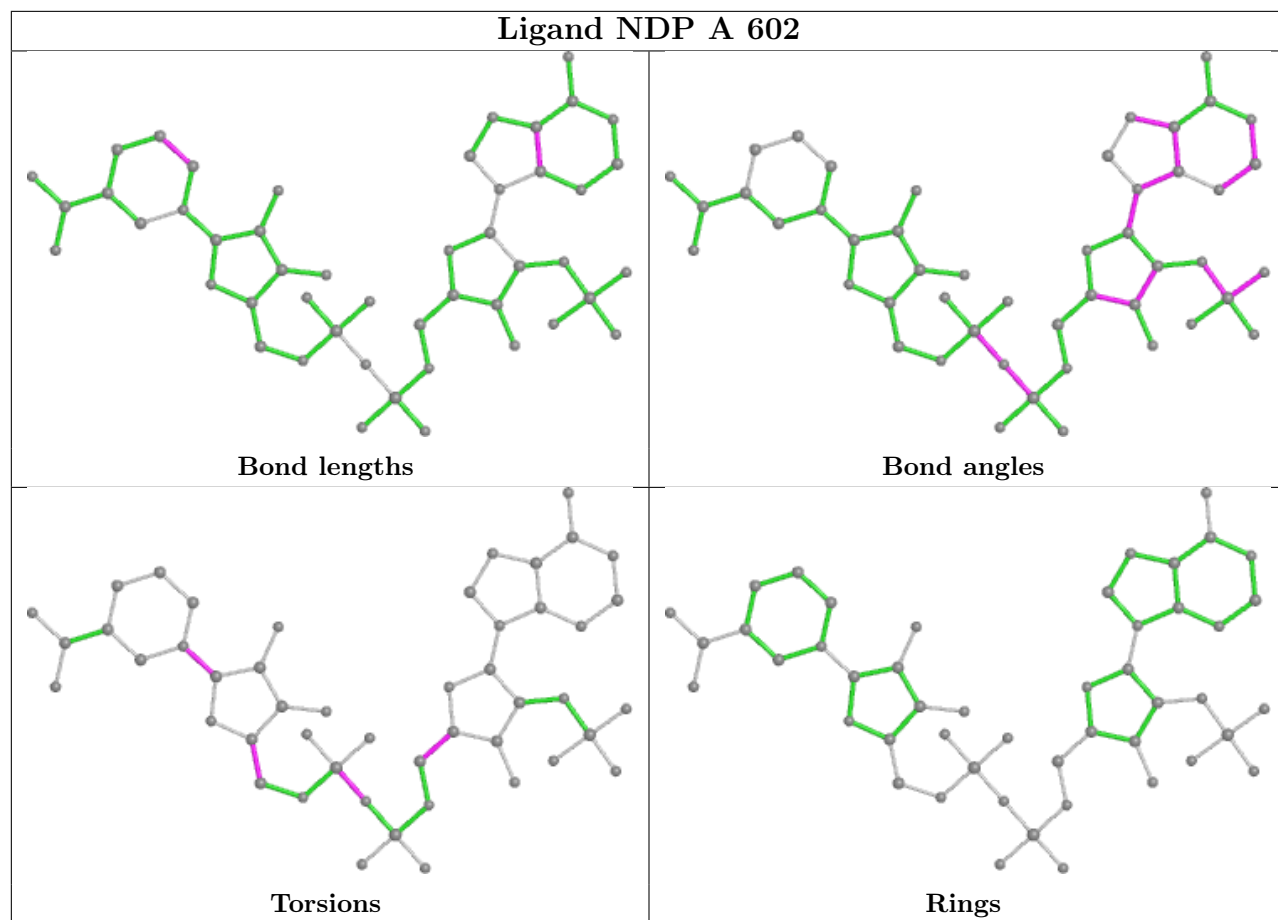


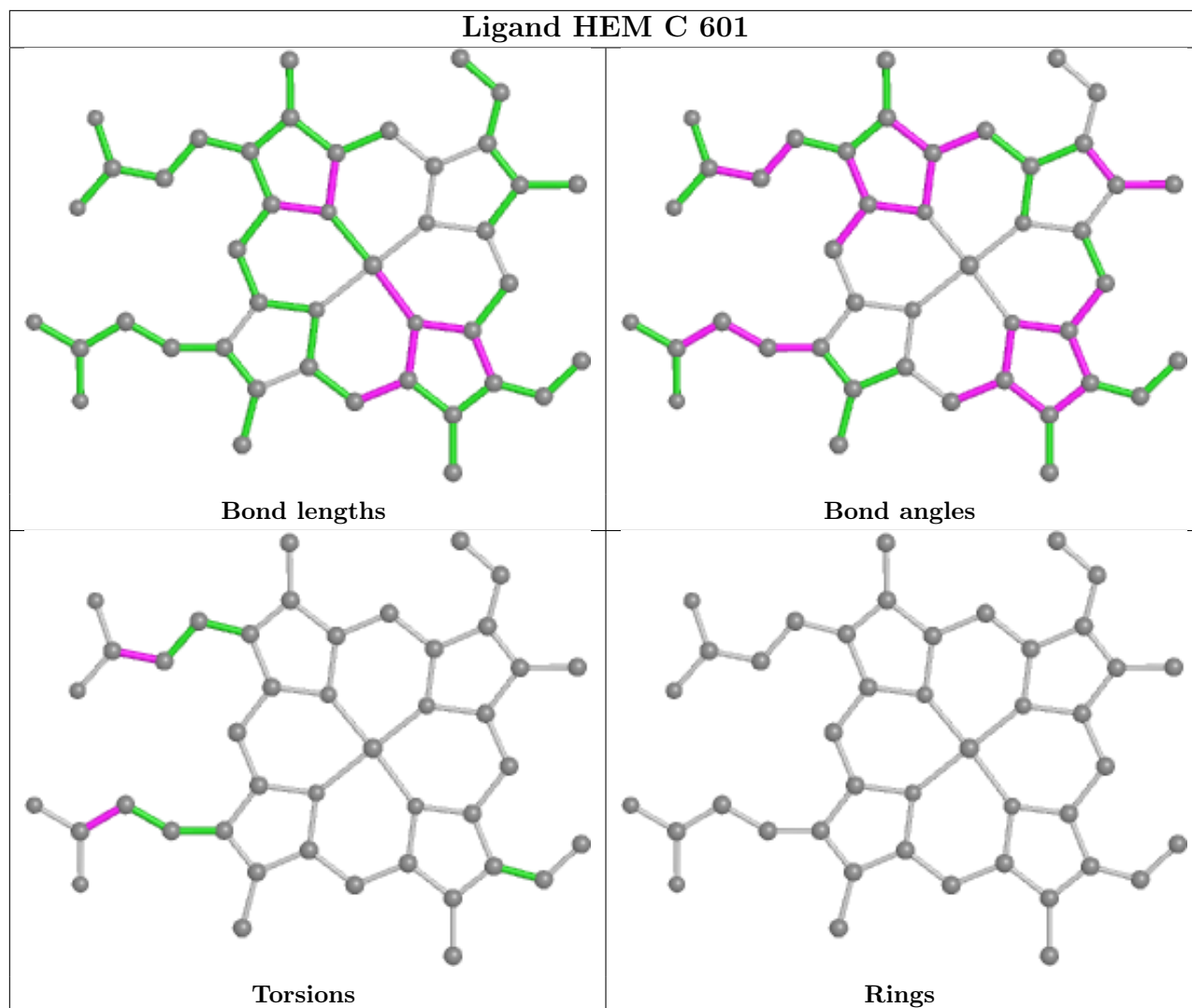












## 5.7 Other polymers [i](#)

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
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	376:CYS	C	377:PRO	N	1.20