



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 22, 2021 – 10:02 am BST

PDB ID : 7O1X  
Title : Unspecific peroxygenase from Hypoxylon sp. EC38 in complex with 1-phenylimidazole  
Authors : Rotilio, L.; Mattevi, A.  
Deposited on : 2021-03-30  
Resolution : 1.60 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.1

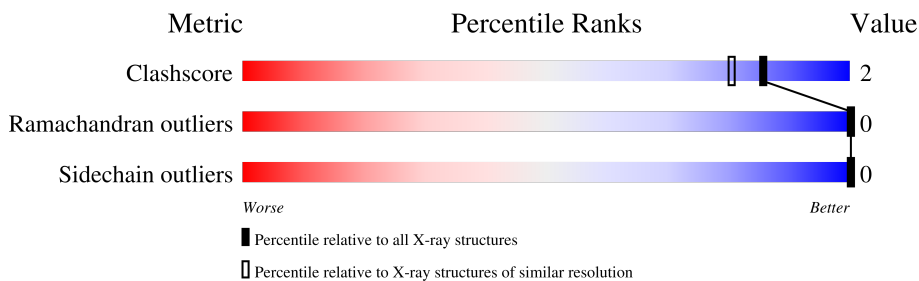
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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))
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	261	 81% 5% 14%

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 2108 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.

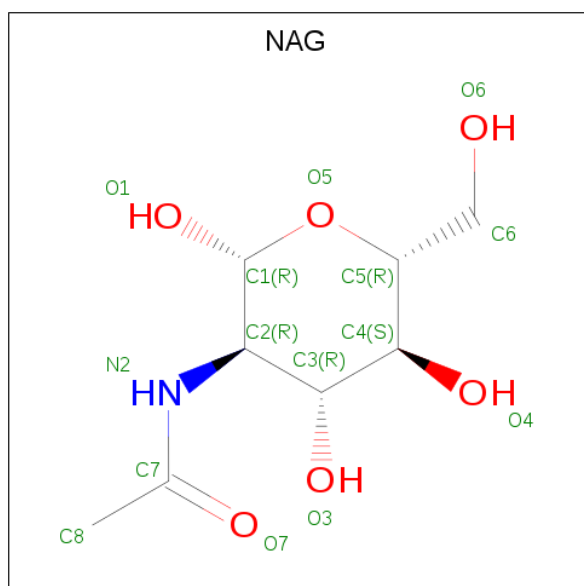
- Molecule 1 is a protein called Peroxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	225	1774	1127	297	346	4	0	6	0

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

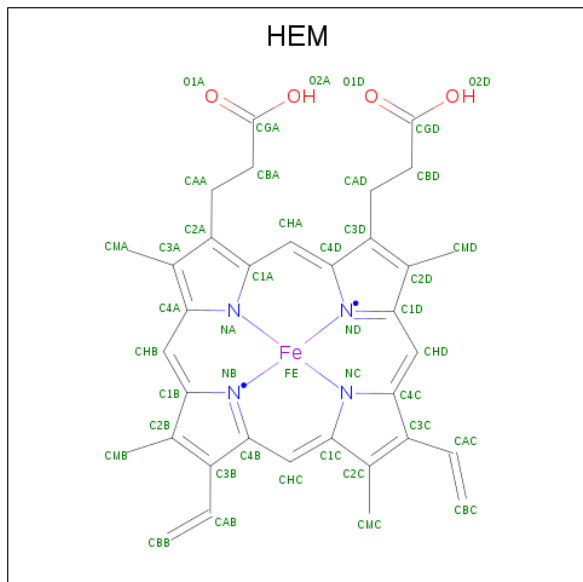
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
2	A	1	1	1	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



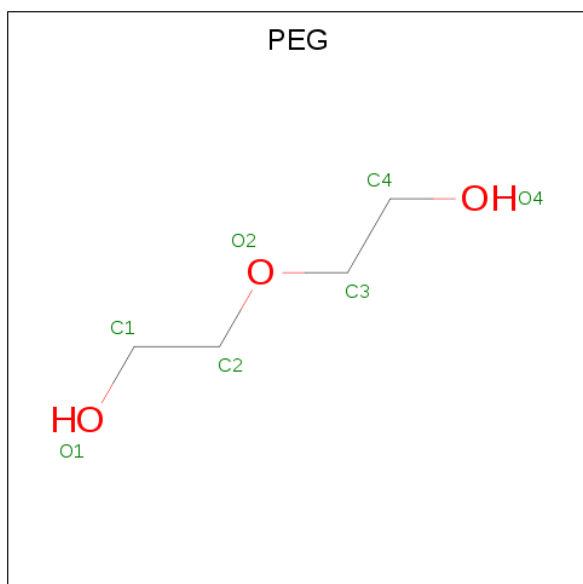
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	14	8	1	5	0	0
3	A	1	14	8	1	5	0	0

- Molecule 4 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



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

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



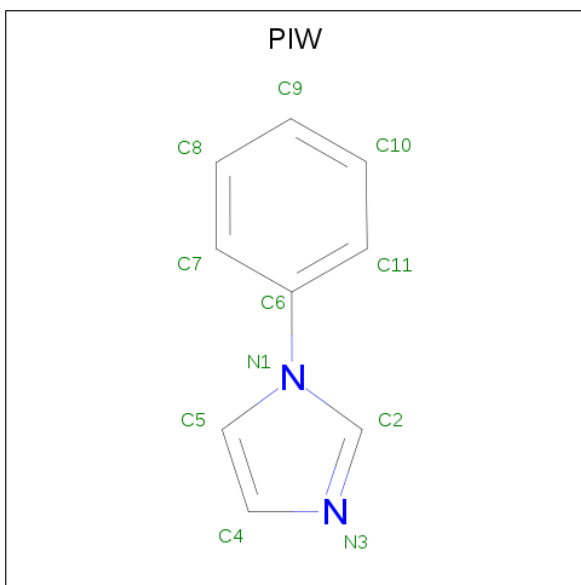
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	7	4	3	0	0

*Continued on next page...*

Continued from previous page...

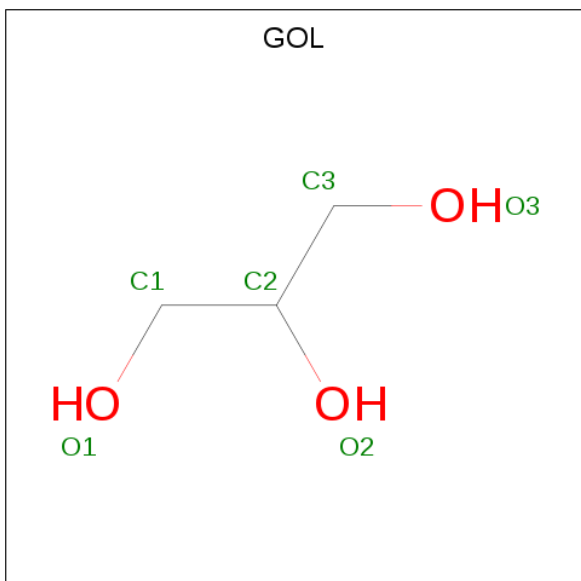
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is 1-phenyl-1H-imidazole (three-letter code: PIW) (formula:  $C_9H_8N_2$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	N	0	0
			11	9	2		

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 8 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	231	Total	O	0	0
			231	231		

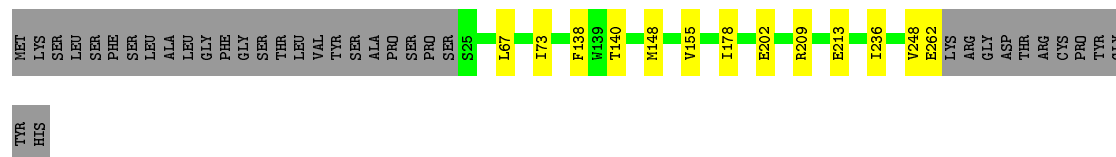
### 3 Residue-property plots [i](#)

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.

Note EDS was not executed.

- Molecule 1: Peroxygenase

Chain A:  81% 5% 14%



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.53Å 71.53Å 153.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.08 – 1.60	Depositor
% Data completeness (in resolution range)	99.9 (48.08-1.60)	Depositor
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.8.0253	Depositor
R, $R_{free}$	0.134 , 0.159	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2108	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP



## 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, GOL, MG, PEG, PIW, NAG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.96	5/1828 (0.3%)	0.95	1/2493 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	202	GLU	CD-OE1	-8.83	1.16	1.25
1	A	262	GLU	CD-OE1	8.71	1.35	1.25
1	A	248	VAL	C-N	7.87	1.52	1.34
1	A	213	GLU	CD-OE1	6.31	1.32	1.25
1	A	209	ARG	NE-CZ	5.30	1.40	1.33

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	262	GLU	CA-C-O	-7.32	104.74	120.10

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	1774	0	1706	7	1
2	A	1	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	28	0	26	0	0
4	A	43	0	30	1	0
5	A	14	0	20	0	0
6	A	11	0	8	0	0
7	A	6	0	8	0	0
8	A	231	0	0	0	0
All	All	2108	0	1798	8	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 2.

All (8) 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:A:140:THR:N	1:A:148[B]:MET:HE3	2.17	0.60
1:A:138:PHE:CE1	1:A:155[B]:VAL:HG11	2.38	0.59
1:A:178:ILE:CD1	1:A:236[A]:ILE:HD11	2.40	0.52
1:A:67:LEU:HB3	1:A:73:ILE:HB	1.97	0.46
1:A:138:PHE:CE2	1:A:155[A]:VAL:HG21	2.54	0.42
1:A:140:THR:H	1:A:148[B]:MET:HE3	1.83	0.42
4:A:303:HEM:HHD	4:A:303:HEM:HBC2	2.01	0.42
1:A:178:ILE:HD13	1:A:236[A]:ILE:HD11	2.02	0.41

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)
1:A:140:THR:OG1	1:A:140:THR:OG1[8_555]	2.13	0.07

## 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	228/261 (87%)	223 (98%)	5 (2%)	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	195/219 (89%)	195 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	28	GLN
1	A	93	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 1 is monoatomic - leaving 7 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	A	302	1	14,14,15	1.35	2 (14%)	17,19,21	1.34	4 (23%)
3	NAG	A	304	1	14,14,15	1.14	1 (7%)	17,19,21	1.96	7 (41%)
5	PEG	A	307	-	6,6,6	0.58	0	5,5,5	0.53	0
7	GOL	A	308	-	5,5,5	0.35	0	5,5,5	0.55	0
5	PEG	A	305	-	6,6,6	0.43	0	5,5,5	0.55	0
6	PIW	A	306	4	12,12,12	2.89	4 (33%)	13,15,15	2.81	8 (61%)
4	HEM	A	303	6,2,1	27,50,50	1.89	5 (18%)	17,82,82	1.78	6 (35%)

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	NAG	A	302	1	-	0/6/23/26	0/1/1/1
3	NAG	A	304	1	-	0/6/23/26	0/1/1/1
5	PEG	A	307	-	-	3/4/4/4	-
7	GOL	A	308	-	-	2/4/4/4	-
5	PEG	A	305	-	-	2/4/4/4	-
6	PIW	A	306	4	-	0/4/4/4	0/2/2/2
4	HEM	A	303	6,2,1	-	0/6/54/54	-

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	306	PIW	C6-N1	-6.58	1.34	1.45
4	A	303	HEM	C3B-C2B	-6.00	1.32	1.40
6	A	306	PIW	C5-N1	-5.67	1.31	1.39
4	A	303	HEM	C3C-C2C	-4.17	1.34	1.40
4	A	303	HEM	C4D-C3D	4.00	1.51	1.42
3	A	304	NAG	C2-N2	-2.96	1.41	1.46
3	A	302	NAG	C1-C2	2.92	1.56	1.52
6	A	306	PIW	C10-C11	2.83	1.44	1.38

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	302	NAG	C4-C5	2.68	1.58	1.53
6	A	306	PIW	C8-C7	2.26	1.43	1.38
4	A	303	HEM	C4A-NA	2.20	1.40	1.36
4	A	303	HEM	C3C-CAC	-2.04	1.43	1.47

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	306	PIW	C5-N1-C2	5.12	117.78	108.50
6	A	306	PIW	C5-N1-C6	-3.86	120.07	126.33
6	A	306	PIW	C4-C5-N1	-3.67	100.02	106.50
3	A	304	NAG	O5-C5-C6	-3.39	101.88	107.20
3	A	304	NAG	O7-C7-C8	3.38	128.33	122.06
4	A	303	HEM	C1D-C2D-C3D	3.32	109.31	107.00
6	A	306	PIW	C11-C6-N1	3.32	123.76	119.33
6	A	306	PIW	C11-C6-C7	-3.24	116.53	121.33
3	A	304	NAG	C8-C7-N2	-3.21	110.67	116.10
6	A	306	PIW	C10-C11-C6	3.05	122.59	118.63
6	A	306	PIW	C5-C4-N3	3.01	112.85	107.57
3	A	304	NAG	O5-C5-C4	-2.98	103.57	110.83
4	A	303	HEM	CMC-C2C-C3C	2.72	129.76	124.68
3	A	302	NAG	C6-C5-C4	2.63	119.16	113.00
4	A	303	HEM	CMD-C2D-C1D	-2.50	124.62	128.46
3	A	302	NAG	C8-C7-N2	-2.46	111.93	116.10
4	A	303	HEM	CMA-C3A-C2A	2.45	129.57	124.94
3	A	304	NAG	C6-C5-C4	2.43	118.69	113.00
4	A	303	HEM	CMA-C3A-C4A	-2.43	124.73	128.46
3	A	302	NAG	O6-C6-C5	-2.37	103.16	111.29
6	A	306	PIW	C8-C7-C6	2.30	121.62	118.63
3	A	302	NAG	O7-C7-C8	2.18	126.11	122.06
3	A	304	NAG	C1-C2-N2	2.16	114.17	110.49
3	A	304	NAG	O3-C3-C2	2.02	113.65	109.47
4	A	303	HEM	C4A-C3A-C2A	-2.02	105.59	107.00

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	308	GOL	C1-C2-C3-O3
5	A	305	PEG	O1-C1-C2-O2
5	A	305	PEG	O2-C3-C4-O4
7	A	308	GOL	O2-C2-C3-O3

*Continued on next page...*

Continued from previous page...

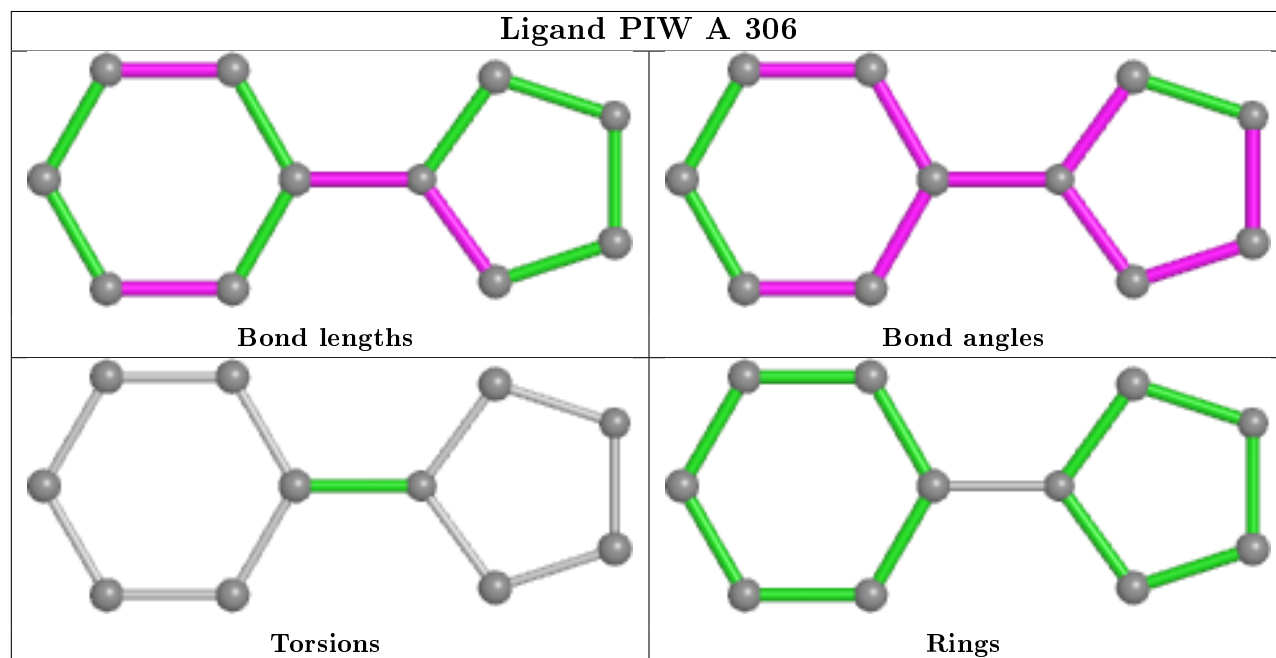
Mol	Chain	Res	Type	Atoms
5	A	307	PEG	O1-C1-C2-O2
5	A	307	PEG	C1-C2-O2-C3
5	A	307	PEG	C4-C3-O2-C2

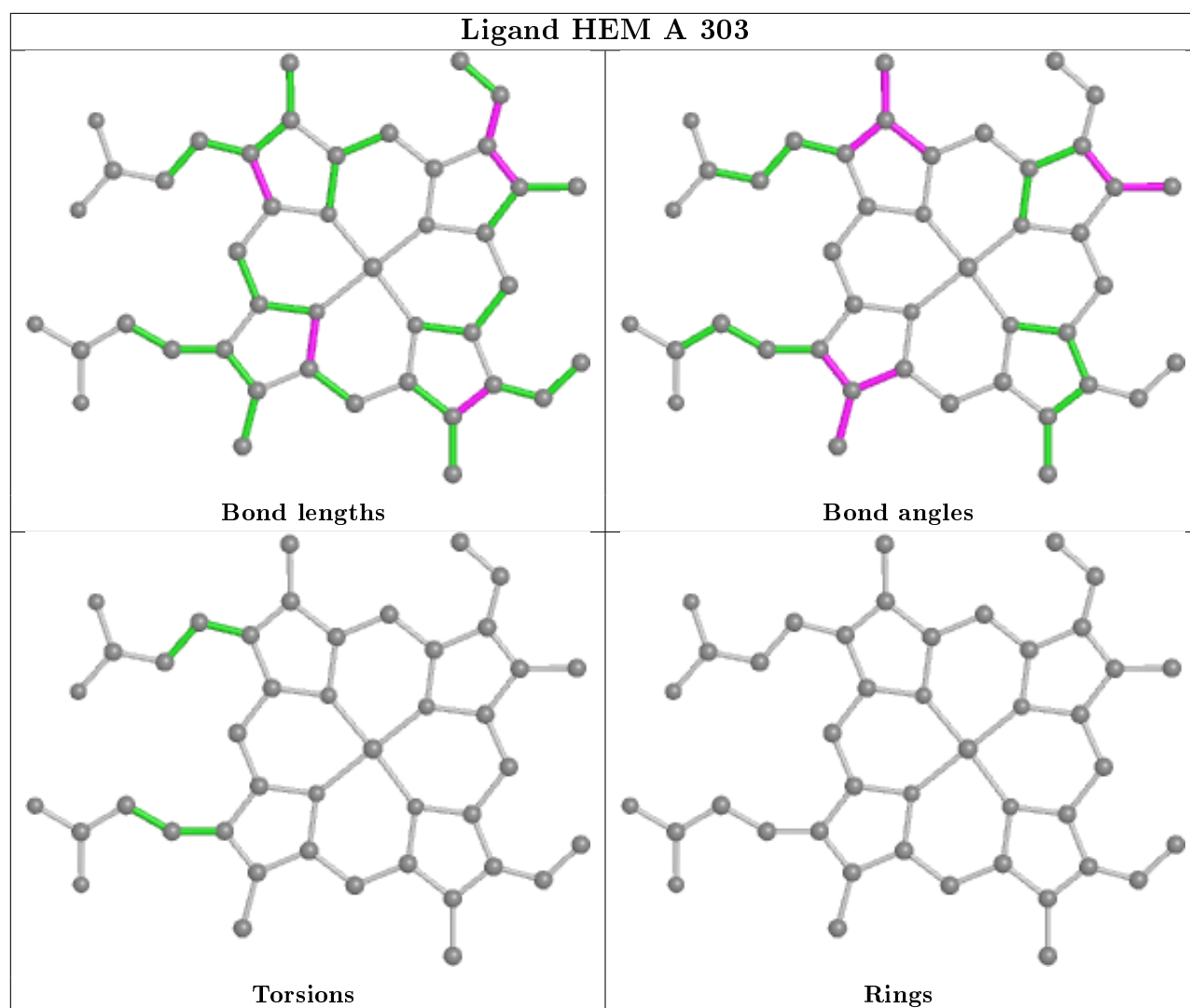
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	303	HEM	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](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.