



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 22, 2021 – 10:02 AM BST

PDB ID : 7O1R
Title : Unspecific peroxygenase from Hypoxylon sp. EC38 in complex with imidazole
Authors : Rotilio, L.; Mattevi, A.
Deposited on : 2021-03-30
Resolution : 1.30 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

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

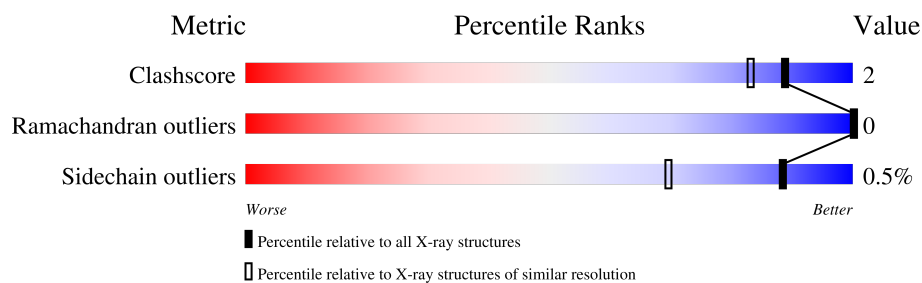
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.30 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	1101 (1.30-1.30)
Ramachandran outliers	138981	1058 (1.30-1.30)
Sidechain outliers	138945	1058 (1.30-1.30)

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 ≥ 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	 84% . 13%

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 2135 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.

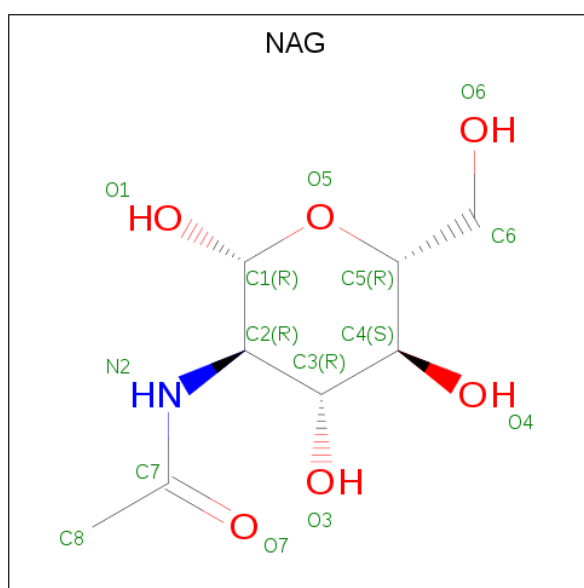
- Molecule 1 is a protein called Peroxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	227	1814	1151	304	356	3	0	8	0

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

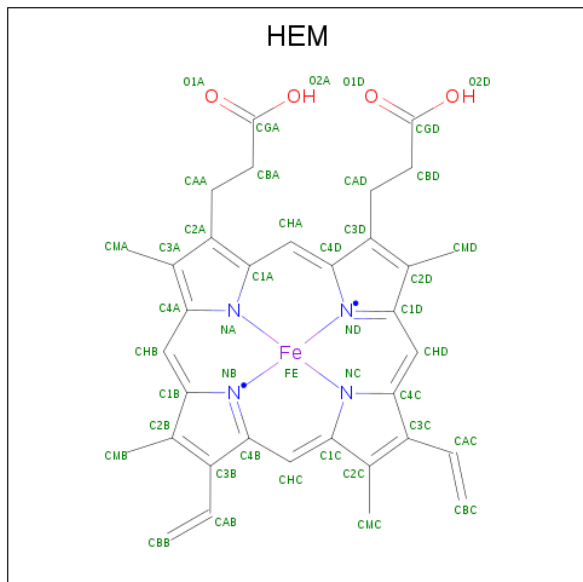
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
2	A	1	1	1	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



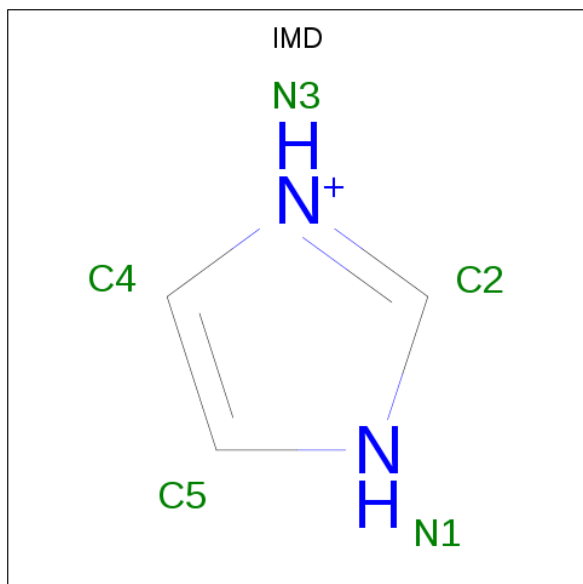
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	14	8	1	5	0	0
3	A	1	14	8	1	5	0	0

- 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 IMIDAZOLE (three-letter code: IMD) (formula: $C_3H_5N_2$) (labeled as "Ligand of Interest" by depositor).



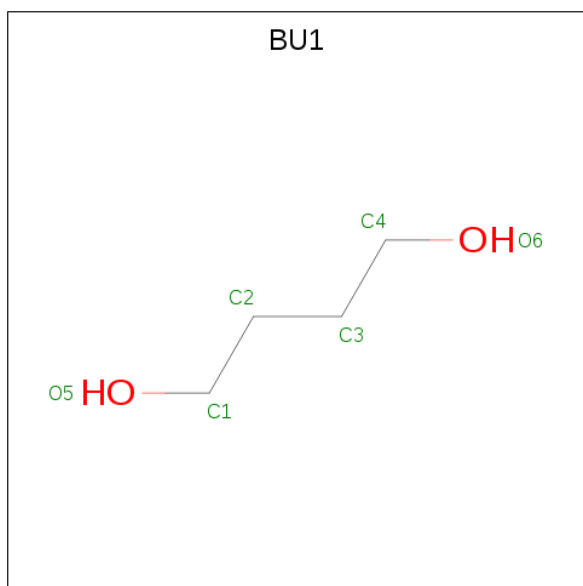
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	N		
5	A	1	5	3	2	0	0

Continued on next page...

Continued from previous page...

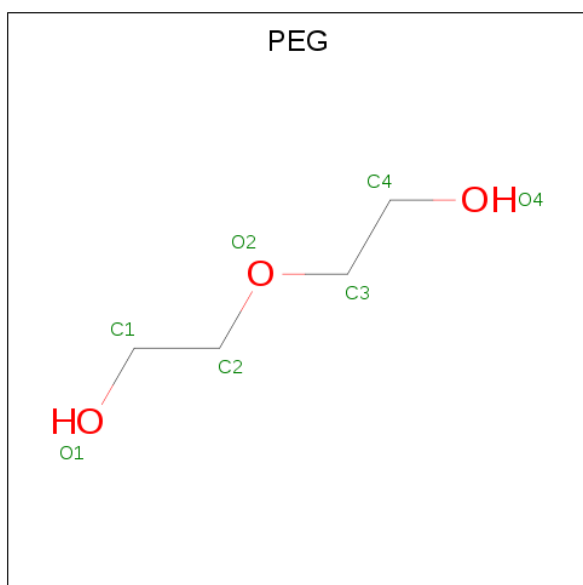
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	N		
5	A	1	5	3	2	0	0

- Molecule 6 is 1,4-BUTANEDIOL (three-letter code: BU1) (formula: C₄H₁₀O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
6	A	1	6	4	2	0	0

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



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

- Molecule 8 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	222	Total	O	0	0
			222	222		

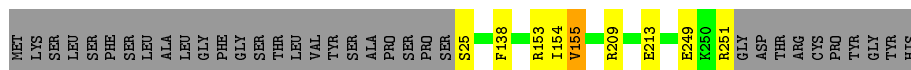
3 Residue-property plots [i](#)

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:  84% 13%



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, α , β , γ	71.73Å 71.73Å 153.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.21 – 1.30	Depositor
% Data completeness (in resolution range)	100.0 (48.21-1.30)	Depositor
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.8.0253	Depositor
R, R_{free}	0.140 , 0.158	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2135	wwPDB-VP
Average B, all atoms (Å ²)	19.0	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: PEG, IMD, MG, HEM, BU1, 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	0.92	3/1874 (0.2%)	1.08	6/2554 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	213	GLU	CD-OE1	9.86	1.36	1.25
1	A	249	GLU	CD-OE2	6.09	1.32	1.25
1	A	213	GLU	CD-OE2	5.44	1.31	1.25

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	209	ARG	NE-CZ-NH1	15.41	128.00	120.30
1	A	251	ARG	CA-C-O	-10.50	98.05	120.10
1	A	209	ARG	NE-CZ-NH2	-8.71	115.95	120.30
1	A	153	ARG	NE-CZ-NH2	7.46	124.03	120.30
1	A	155	VAL	CG1-CB-CG2	-6.67	100.24	110.90
1	A	155	VAL	CA-CB-CG1	5.12	118.58	110.90

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	1814	0	1748	2	0
2	A	1	0	0	0	0
3	A	28	0	26	0	0
4	A	43	0	30	2	0
5	A	10	0	9	3	0
6	A	6	0	10	0	0
7	A	11	0	12	0	0
8	A	222	0	0	4	0
All	All	2135	0	1835	7	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 2.

All (7) 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)
5:A:308:IMD:H5	8:A:471:HOH:O	0.79	0.96
1:A:138:PHE:CE1	1:A:155:VAL:HG11	2.39	0.58
5:A:308:IMD:C5	8:A:471:HOH:O	1.69	0.49
1:A:25:SER:N	8:A:404:HOH:O	2.49	0.44
4:A:303:HEM:HBB2	4:A:303:HEM:HMB2	2.01	0.42
5:A:308:IMD:C2	8:A:452:HOH:O	2.68	0.42
4:A:303:HEM:HBC2	4:A:303:HEM:HHD	2.02	0.40

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 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	234/261 (90%)	229 (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	200/219 (91%)	199 (100%)	1 (0%)	88 69

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	154	ILE

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	91	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 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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	305	1	14,14,15	1.06	1 (7%)	17,19,21	1.86	5 (29%)
5	IMD	A	304	4	3,5,5	0.38	0	4,5,5	0.42	0
7	PEG	A	309	-	3,3,6	0.62	0	2,2,5	0.86	0
6	BU1	A	306	-	5,5,5	0.92	0	4,4,4	0.70	0
3	NAG	A	302	1	14,14,15	1.14	2 (14%)	17,19,21	1.30	2 (11%)
7	PEG	A	307	-	6,6,6	0.34	0	5,5,5	0.32	0
4	HEM	A	303	5,1,2	27,50,50	1.66	5 (18%)	17,82,82	2.28	5 (29%)
5	IMD	A	308	-	3,5,5	0.88	0	4,5,5	0.88	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
3	NAG	A	305	1	-	0/6/23/26	0/1/1/1
7	PEG	A	309	-	-	0/0/1/4	-
5	IMD	A	304	4	-	-	0/1/1/1
6	BU1	A	306	-	-	2/3/3/3	-
3	NAG	A	302	1	-	1/6/23/26	0/1/1/1
7	PEG	A	307	-	-	3/4/4/4	-
4	HEM	A	303	5,1,2	-	0/6/54/54	-
5	IMD	A	308	-	-	-	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	303	HEM	C3B-C2B	-4.63	1.34	1.40
4	A	303	HEM	C3C-C2C	-3.29	1.35	1.40
4	A	303	HEM	C4D-C3D	3.16	1.49	1.42
3	A	302	NAG	O6-C6	2.86	1.54	1.42
4	A	303	HEM	C4B-CHC	-2.81	1.33	1.41
4	A	303	HEM	CAD-C3D	-2.67	1.47	1.52
3	A	302	NAG	O5-C5	2.30	1.48	1.43
3	A	305	NAG	C3-C2	2.16	1.57	1.52

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	303	HEM	CAD-CBD-CGD	6.94	124.32	112.67
3	A	305	NAG	C6-C5-C4	4.14	122.69	113.00
3	A	302	NAG	O6-C6-C5	-3.36	99.75	111.29
3	A	302	NAG	C1-O5-C5	3.27	116.62	112.19
4	A	303	HEM	CAA-CBA-CGA	-2.94	107.74	112.67
3	A	305	NAG	O7-C7-C8	2.92	127.48	122.06
3	A	305	NAG	C8-C7-N2	-2.86	111.26	116.10
3	A	305	NAG	O6-C6-C5	-2.57	102.46	111.29
4	A	303	HEM	C3B-C4B-NB	-2.55	105.92	109.21
4	A	303	HEM	CBD-CAD-C3D	2.48	117.04	112.48
3	A	305	NAG	O5-C1-C2	-2.17	107.86	111.29
4	A	303	HEM	CMD-C2D-C1D	-2.07	125.28	128.46

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	307	PEG	O1-C1-C2-O2
7	A	307	PEG	O2-C3-C4-O4
6	A	306	BU1	O5-C1-C2-C3
6	A	306	BU1	C1-C2-C3-C4
3	A	302	NAG	C4-C5-C6-O6
7	A	307	PEG	C4-C3-O2-C2

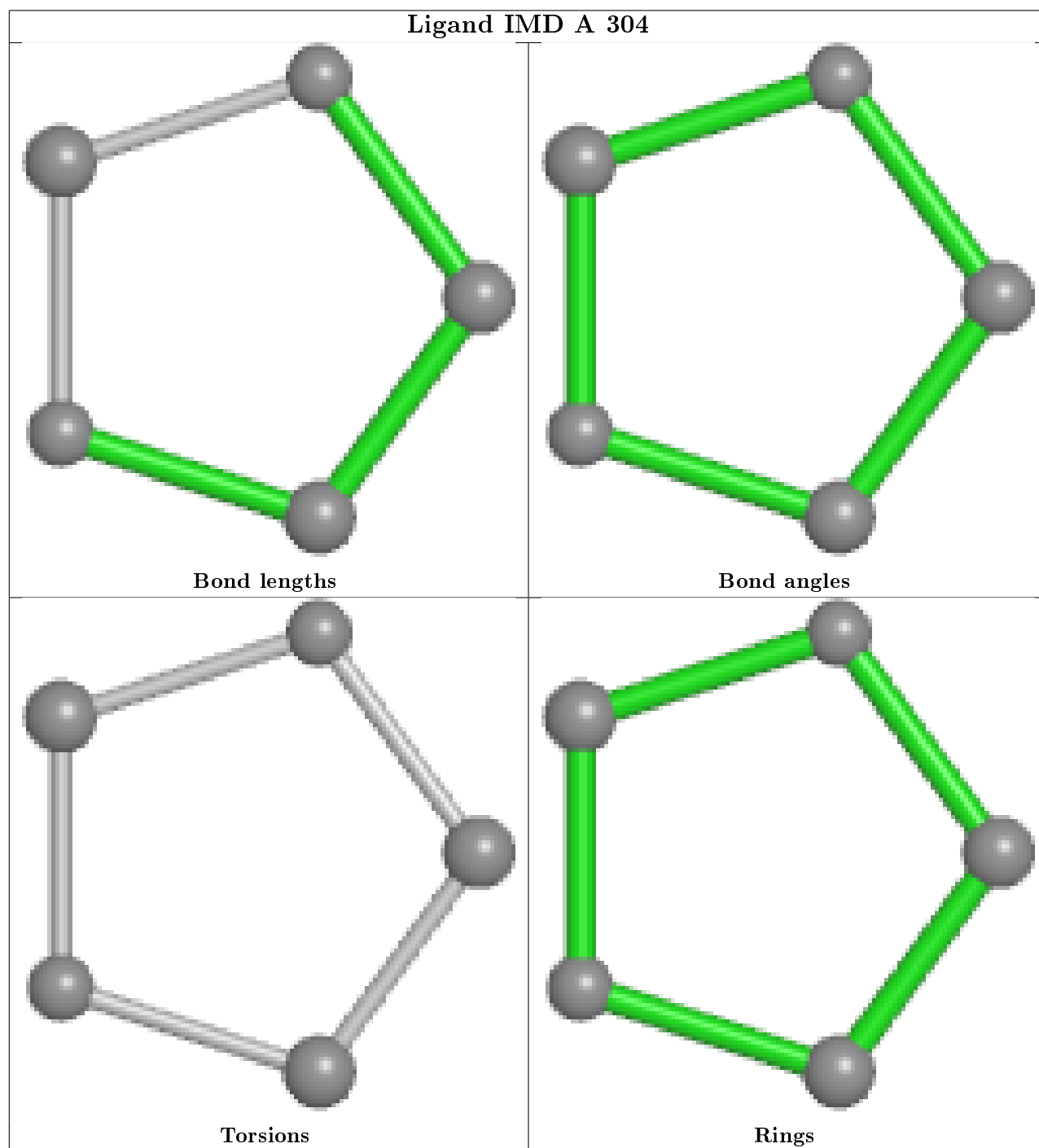
There are no ring outliers.

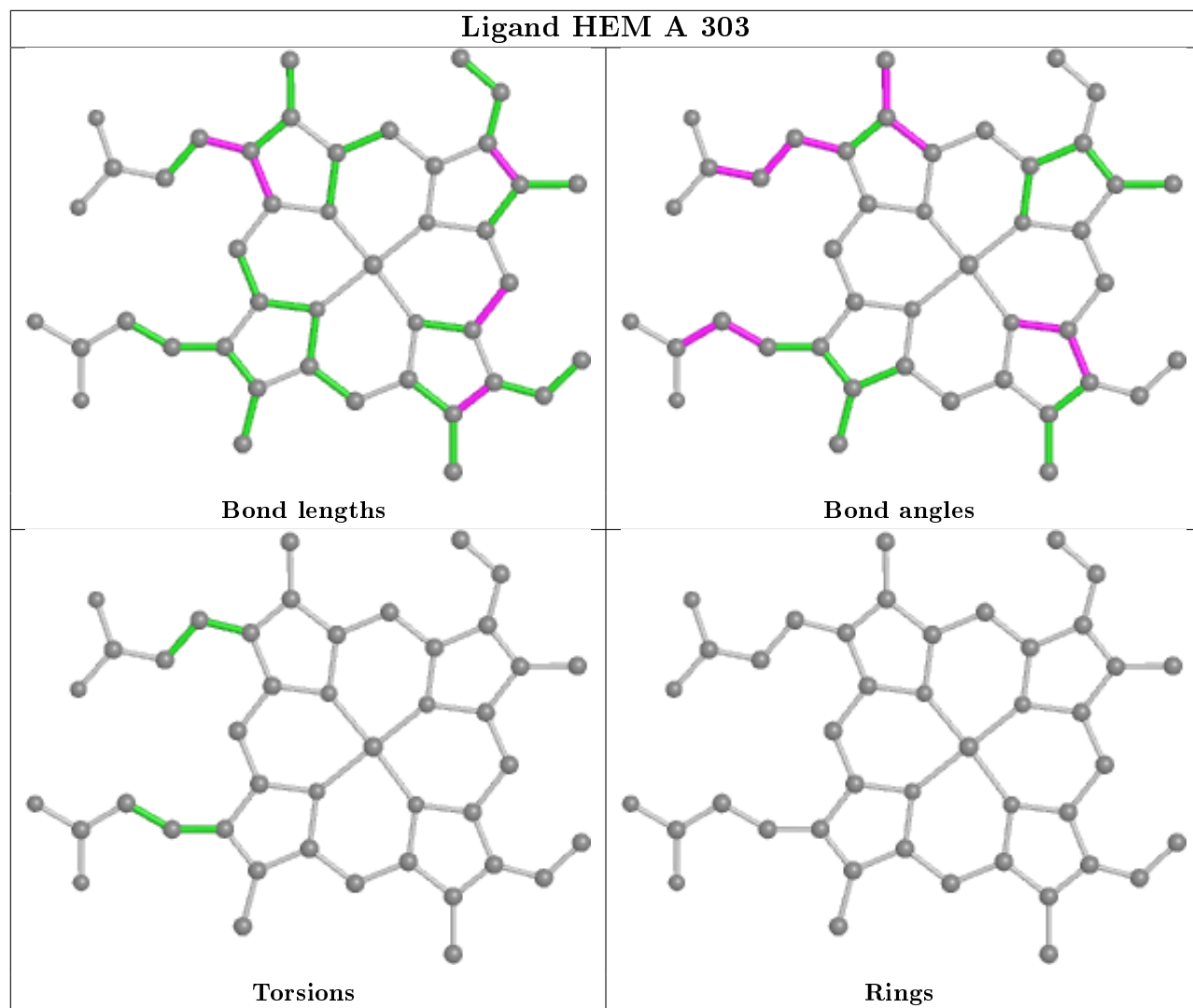
2 monomers are involved in 5 short contacts:

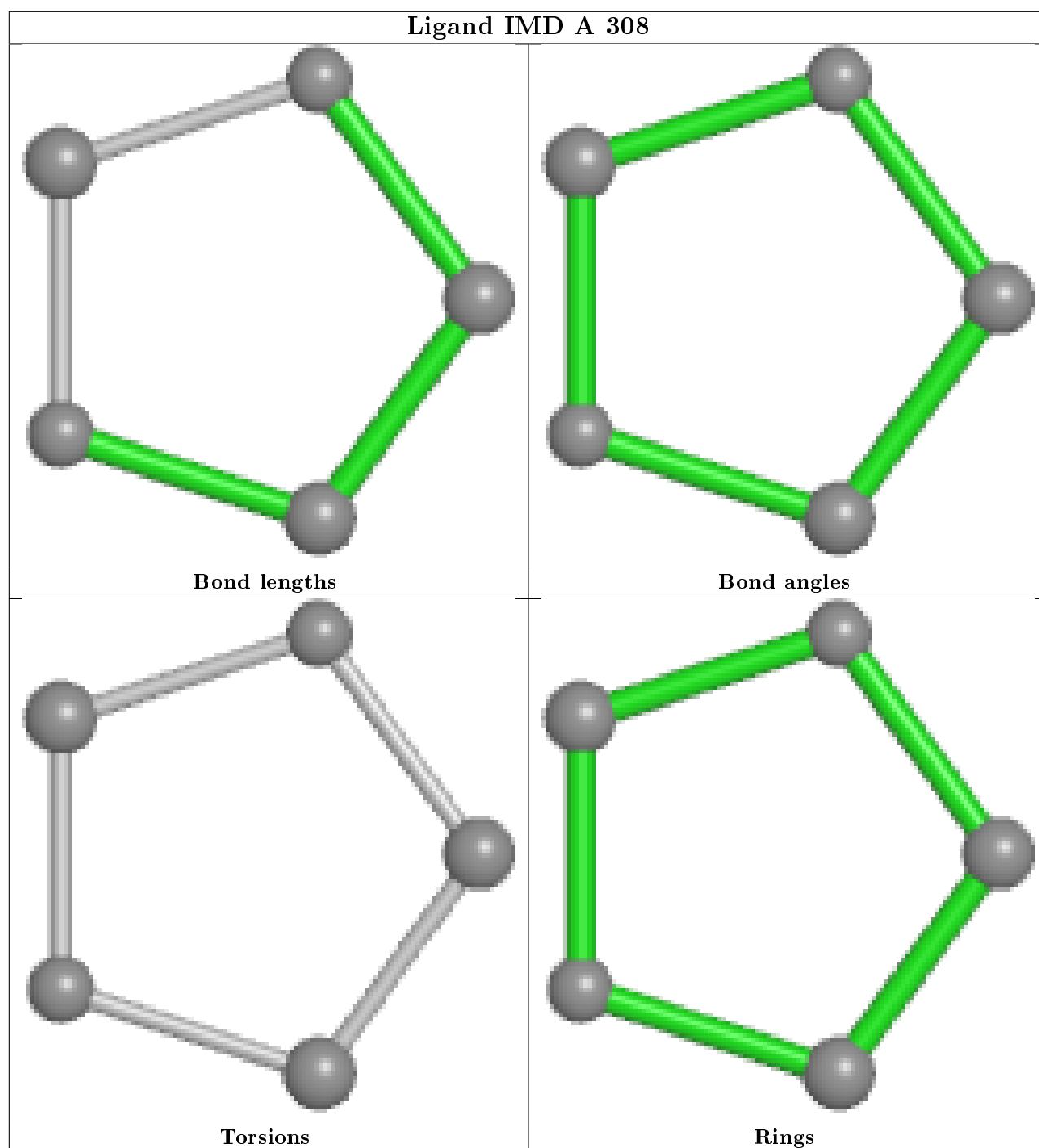
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	303	HEM	2	0
5	A	308	IMD	3	0

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.