



## wwPDB EM Validation Summary Report ⓘ

Nov 29, 2022 – 03:39 AM EST

PDB ID : 7UTE  
EMDB ID : EMD-26768  
Title : MicroED structure of Aeropyrum pernix protoglobin mutant  
Authors : Danelius, E.; Gonen, T.; Unge, J.T.  
Deposited on : 2022-04-26  
Resolution : 2.10 Å(reported)

This is a wwPDB EM 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/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:

EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	<b>FAILED</b>
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

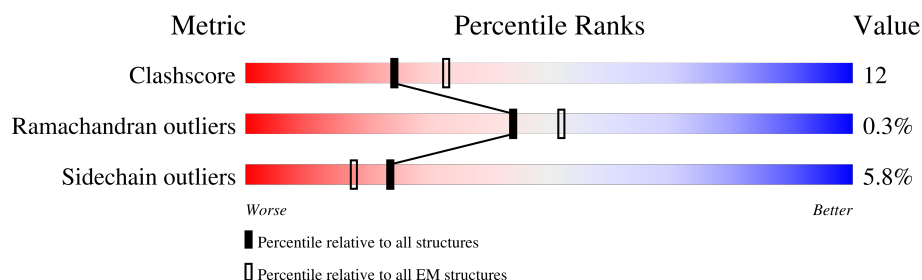
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON CRYSTALLOGRAPHY*

The reported resolution of this entry is 2.10 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)
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	195	69% 24% 5% .
1	B	195	66% 26% . 6%
1	C	195	72% 20% . 6%
1	D	195	71% 21% . 6%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6547 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 Protogloblin ApPgb.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	190	Total	C	N	O	S	1	0
			1577	1022	265	287	3		
1	B	184	Total	C	N	O	S	0	0
			1517	985	252	277	3		
1	C	184	Total	C	N	O	S	0	0
			1517	985	252	277	3		
1	D	184	Total	C	N	O	S	0	0
			1504	976	248	277	3		

There are 28 discrepancies between the modelled and reference sequences:

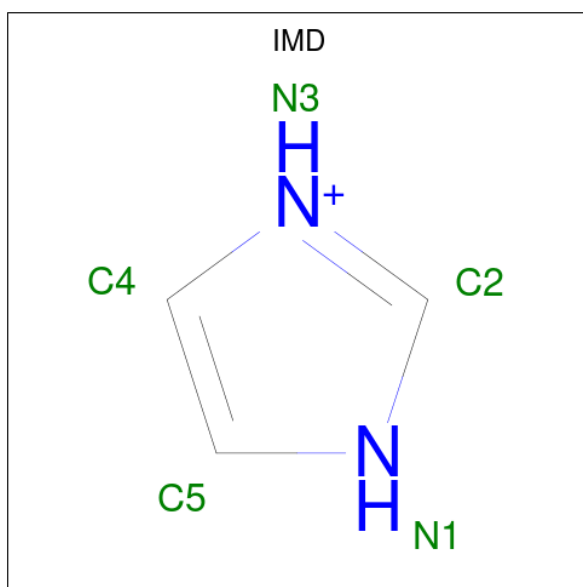
Chain	Residue	Modelled	Actual	Comment	Reference
A	45	GLY	CYS	conflict	UNP Q9YFF4
A	59	LEU	TRP	conflict	UNP Q9YFF4
A	60	VAL	TYR	conflict	UNP Q9YFF4
A	63	ARG	VAL	conflict	UNP Q9YFF4
A	102	SER	CYS	conflict	UNP Q9YFF4
A	145	GLN	PHE	conflict	UNP Q9YFF4
A	149	LEU	ILE	conflict	UNP Q9YFF4
B	45	GLY	CYS	conflict	UNP Q9YFF4
B	59	LEU	TRP	conflict	UNP Q9YFF4
B	60	VAL	TYR	conflict	UNP Q9YFF4
B	63	ARG	VAL	conflict	UNP Q9YFF4
B	102	SER	CYS	conflict	UNP Q9YFF4
B	145	GLN	PHE	conflict	UNP Q9YFF4
B	149	LEU	ILE	conflict	UNP Q9YFF4
C	45	GLY	CYS	conflict	UNP Q9YFF4
C	59	LEU	TRP	conflict	UNP Q9YFF4
C	60	VAL	TYR	conflict	UNP Q9YFF4
C	63	ARG	VAL	conflict	UNP Q9YFF4
C	102	SER	CYS	conflict	UNP Q9YFF4
C	145	GLN	PHE	conflict	UNP Q9YFF4
C	149	LEU	ILE	conflict	UNP Q9YFF4
D	45	GLY	CYS	conflict	UNP Q9YFF4

*Continued on next page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	59	LEU	TRP	conflict	UNP Q9YFF4
D	60	VAL	TYR	conflict	UNP Q9YFF4
D	63	ARG	VAL	conflict	UNP Q9YFF4
D	102	SER	CYS	conflict	UNP Q9YFF4
D	145	GLN	PHE	conflict	UNP Q9YFF4
D	149	LEU	ILE	conflict	UNP Q9YFF4

- # HEM

- Molecule 3 is IMIDAZOLE (three-letter code: IMD) (formula:  $\text{C}_3\text{H}_5\text{N}_2$ ).



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

- Molecule 4 is water.

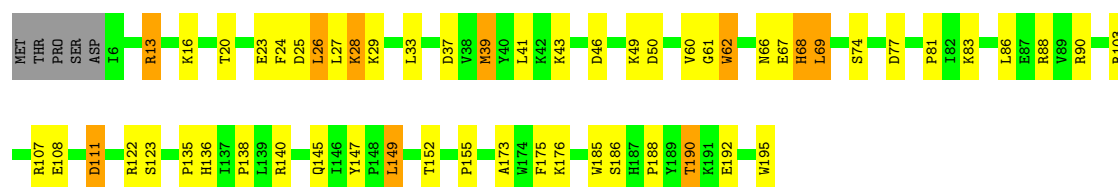
Mol	Chain	Residues	Atoms		AltConf
4	A	58	Total	O	0
			58	58	
4	B	59	Total	O	0
			59	59	
4	C	49	Total	O	0
			49	49	
4	D	74	Total	O	0
			74	74	

### 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.

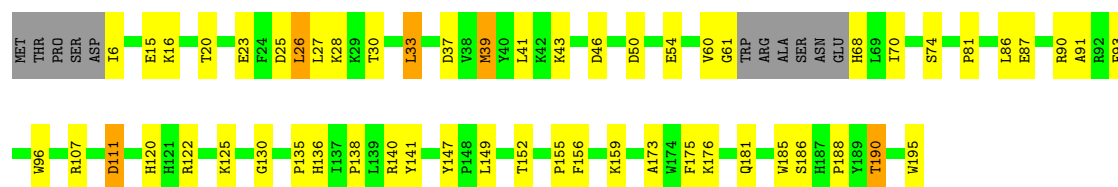
#### • Molecule 1: Protogloblin ApPgb

Chain A: 



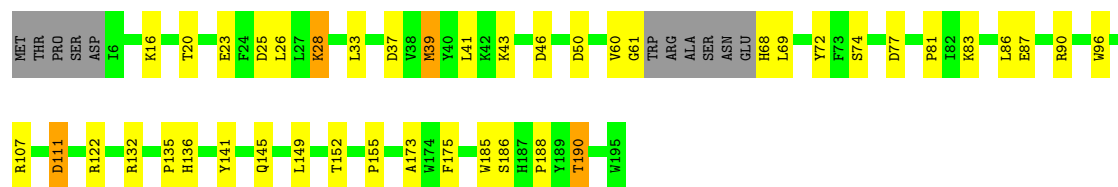
#### • Molecule 1: Protogloblin ApPgb

Chain B: 



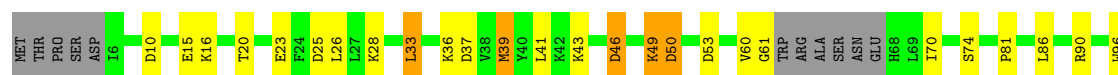
#### • Molecule 1: Protogloblin ApPgb

Chain C: 



#### • Molecule 1: Protogloblin ApPgb

Chain D: 



R107	D111	H120	H121	R122	G130	P135	H136	Y141	T152	P155	A173	W174	F175	W185	S186	H187	P188	Y189	T190	K191	E192	N193	D194	W195
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

## 4 Experimental information

Property	Value	Source
EM reconstruction method	CRYSTALLOGRAPHY	Depositor
Imposed symmetry	3D CRYSTAL, $a=46.2$ Å, $b=58.3$ Å, $c=80.7$ Å, $\alpha=104.1^\circ$ , $\beta=98.6^\circ$ , $\gamma=90.1^\circ$ , space group=P1	Depositor
Number of images used	Not provided	
Resolution determination method	DIFFRACTION PATTERN/LAYERLINES	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	0.00025	Depositor
Minimum defocus (nm)	0	Depositor
Maximum defocus (nm)	0	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor



## 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: IMD, HEM

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	0.57	0/1625	0.84	4/2205 (0.2%)
1	B	0.60	0/1558	0.85	4/2112 (0.2%)
1	C	0.57	0/1558	0.84	4/2112 (0.2%)
1	D	0.56	0/1545	0.83	4/2097 (0.2%)
All	All	0.58	0/6286	0.84	16/8526 (0.2%)

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	26	LEU	CA-CB-CG	-5.91	101.70	115.30
1	C	26	LEU	CA-CB-CG	-5.88	101.76	115.30
1	A	26	LEU	CA-CB-CG	-5.88	101.78	115.30
1	B	26	LEU	CA-CB-CG	-5.87	101.81	115.30
1	C	43	LYS	CB-CG-CD	5.82	126.74	111.60

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	1577	0	1559	43	0
1	B	1517	0	1506	44	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1517	0	1506	35	0
1	D	1504	0	1475	31	0
2	A	43	0	30	2	0
2	B	43	0	30	10	0
2	C	43	0	30	12	0
2	D	43	0	30	7	0
3	A	5	0	4	0	0
3	B	5	0	4	0	0
3	C	5	0	4	0	0
3	D	5	0	4	0	0
4	A	58	0	0	8	0
4	B	59	0	0	9	0
4	C	49	0	0	8	0
4	D	74	0	0	7	0
All	All	6547	0	6182	147	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 12.

The worst 5 of 147 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:145:GLN:HG3	1:C:149:LEU:HD12	1.39	1.04
1:B:181:GLN:OE1	4:B:301:HOH:O	1.82	0.96
1:A:13:ARG:NH2	4:A:302:HOH:O	2.05	0.89
1:D:10:ASP:OD2	4:D:301:HOH:O	1.93	0.87
1:A:192:GLU:OE1	4:A:301:HOH:O	1.93	0.86

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	189/195 (97%)	177 (94%)	9 (5%)	3 (2%)	9	5
1	B	180/195 (92%)	174 (97%)	6 (3%)	0	100	100
1	C	180/195 (92%)	173 (96%)	7 (4%)	0	100	100
1	D	180/195 (92%)	174 (97%)	6 (3%)	0	100	100
All	All	729/780 (94%)	698 (96%)	28 (4%)	3 (0%)	44	32

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	68[A]	HIS
1	A	68[B]	HIS
1	A	69	LEU

### 5.3.2 Protein sidechains ⓘ

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	169/173 (98%)	158 (94%)	11 (6%)	17	14
1	B	163/173 (94%)	154 (94%)	9 (6%)	21	19
1	C	163/173 (94%)	156 (96%)	7 (4%)	29	29
1	D	160/173 (92%)	149 (93%)	11 (7%)	15	12
All	All	655/692 (95%)	617 (94%)	38 (6%)	24	17

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	28	LYS
1	D	70	ILE
1	D	33	LEU
1	D	46	ASP
1	D	175	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	51	GLN
1	B	51	GLN
1	B	68	HIS
1	C	51	GLN
1	D	51	GLN

### 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	IMD	A	202	2	3,5,5	0.41	0	4,5,5	1.02	0
2	HEM	D	201	1,3	41,50,50	1.51	4 (9%)	45,82,82	2.13	15 (33%)
3	IMD	C	202	2	3,5,5	0.41	0	4,5,5	1.02	0
2	HEM	C	201	1,3	41,50,50	1.59	6 (14%)	45,82,82	1.82	10 (22%)
3	IMD	D	202	2	3,5,5	0.41	0	4,5,5	1.02	0
2	HEM	B	201	1,3	41,50,50	1.52	4 (9%)	45,82,82	1.84	8 (17%)
2	HEM	A	201	1,3	41,50,50	1.67	8 (19%)	45,82,82	1.69	10 (22%)
3	IMD	B	202	2	3,5,5	0.41	0	4,5,5	1.02	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	IMD	A	202	2	-	-	0/1/1/1
2	HEM	D	201	1,3	-	4/12/54/54	-
3	IMD	C	202	2	-	-	0/1/1/1
2	HEM	C	201	1,3	-	4/12/54/54	-
3	IMD	D	202	2	-	-	0/1/1/1
2	HEM	B	201	1,3	-	4/12/54/54	-
2	HEM	A	201	1,3	-	4/12/54/54	-
3	IMD	B	202	2	-	-	0/1/1/1

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	201	HEM	C3C-C2C	-6.13	1.31	1.40
2	C	201	HEM	C3C-C2C	-5.87	1.32	1.40
2	A	201	HEM	C3C-C2C	-5.11	1.33	1.40
2	B	201	HEM	C3C-C2C	-5.06	1.33	1.40
2	A	201	HEM	FE-NB	3.73	2.15	1.96

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	201	HEM	CBA-CAA-C2A	-7.56	99.72	112.62
2	B	201	HEM	CBA-CAA-C2A	-5.37	103.46	112.62
2	B	201	HEM	C2C-C3C-C4C	5.31	110.60	106.90
2	C	201	HEM	CBA-CAA-C2A	-5.07	103.96	112.62
2	D	201	HEM	C2C-C3C-C4C	4.86	110.29	106.90

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

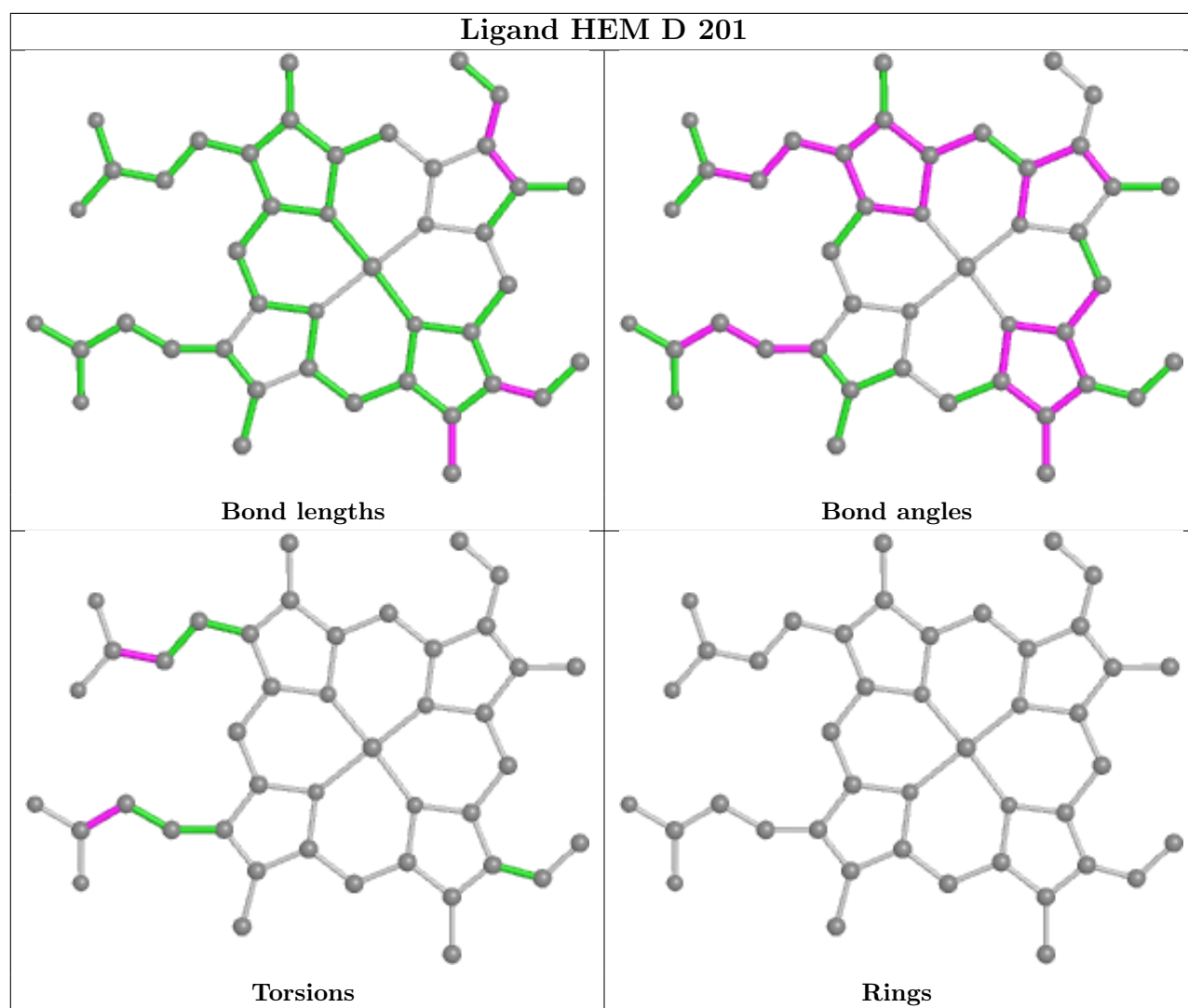
Mol	Chain	Res	Type	Atoms
2	C	201	HEM	CAA-CBA-CGA-O2A
2	B	201	HEM	CAD-CBD-CGD-O1D
2	C	201	HEM	CAD-CBD-CGD-O1D
2	A	201	HEM	CAD-CBD-CGD-O1D
2	B	201	HEM	CAD-CBD-CGD-O2D

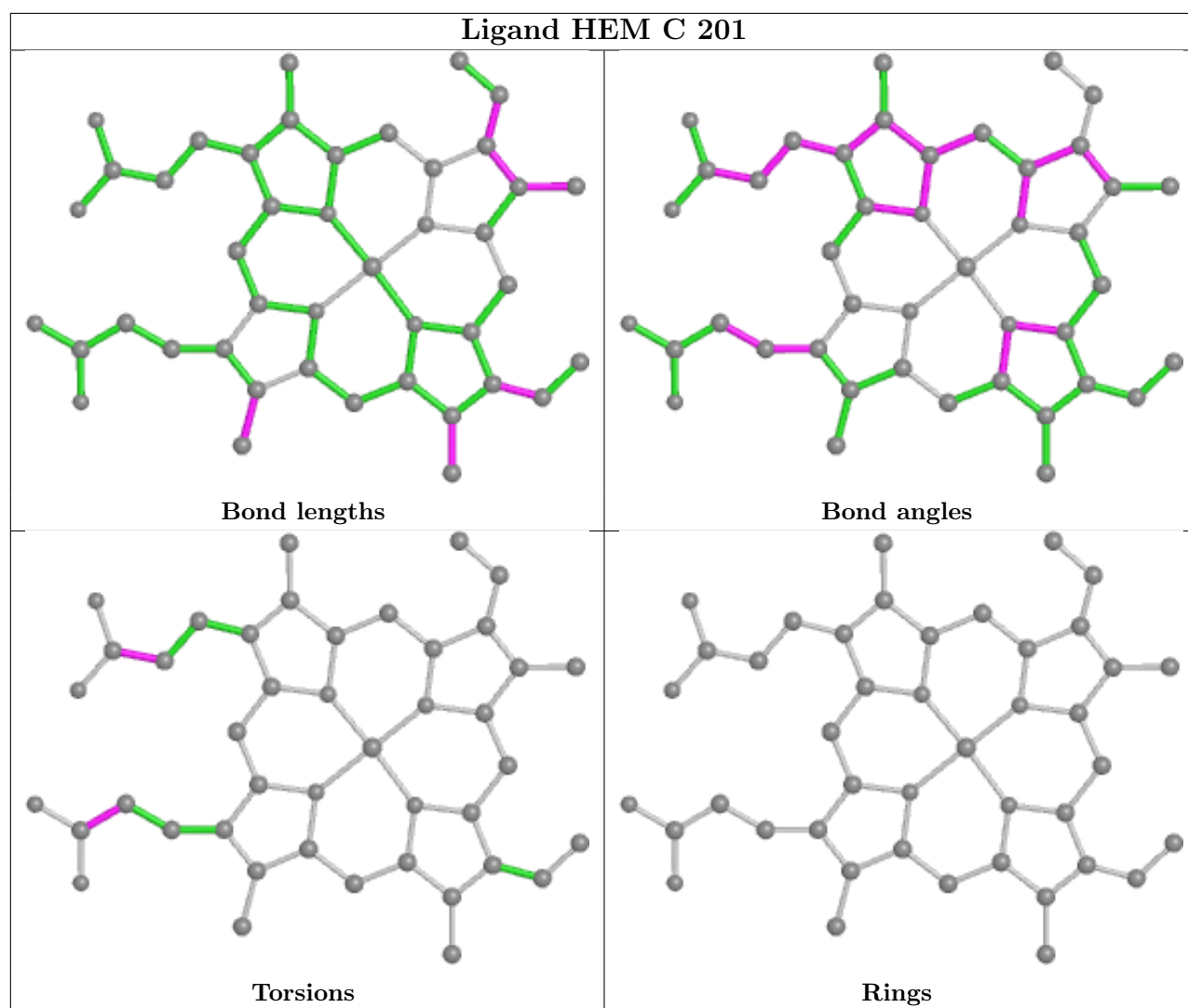
There are no ring outliers.

4 monomers are involved in 31 short contacts:

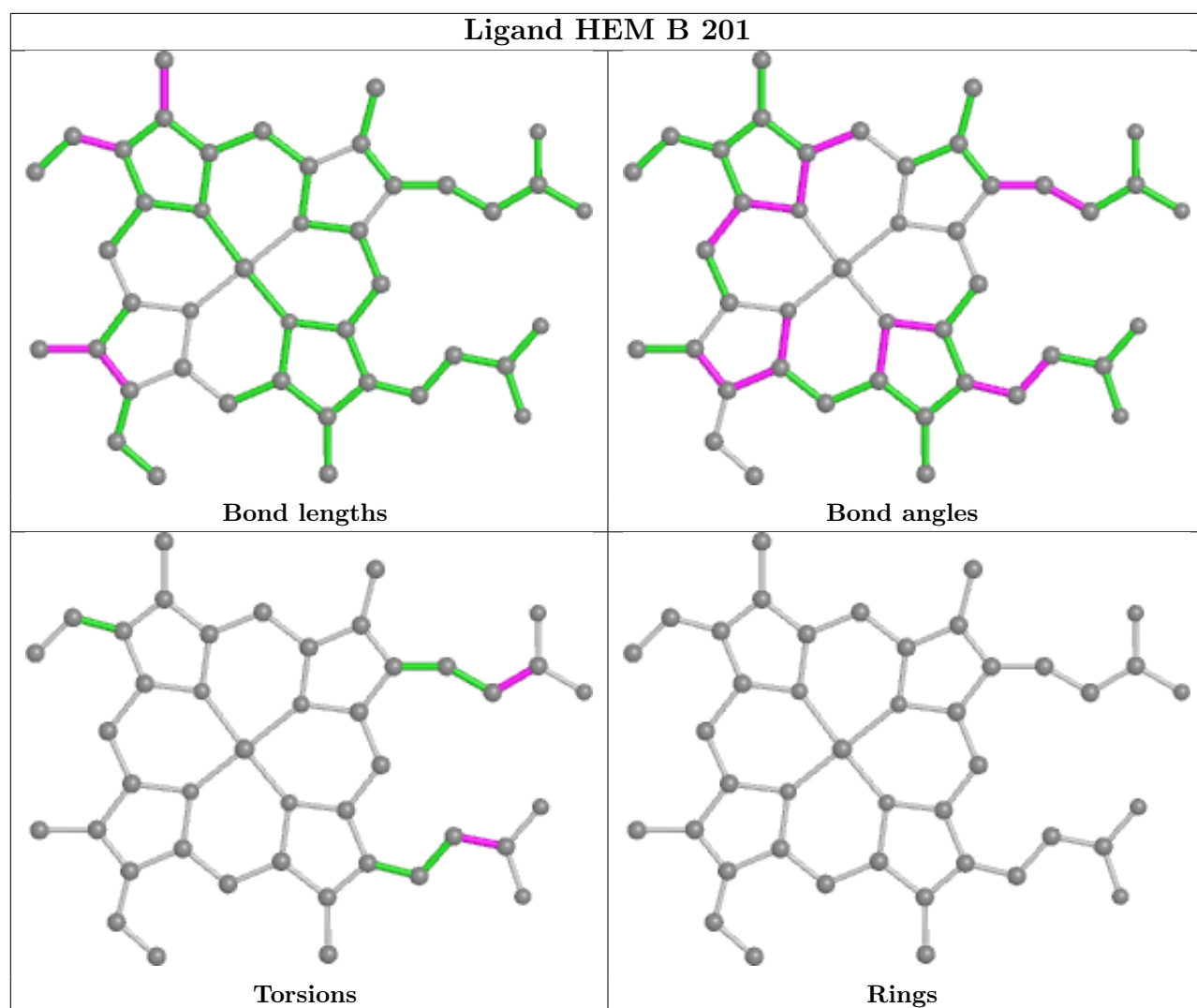
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	201	HEM	7	0
2	C	201	HEM	12	0
2	B	201	HEM	10	0
2	A	201	HEM	2	0

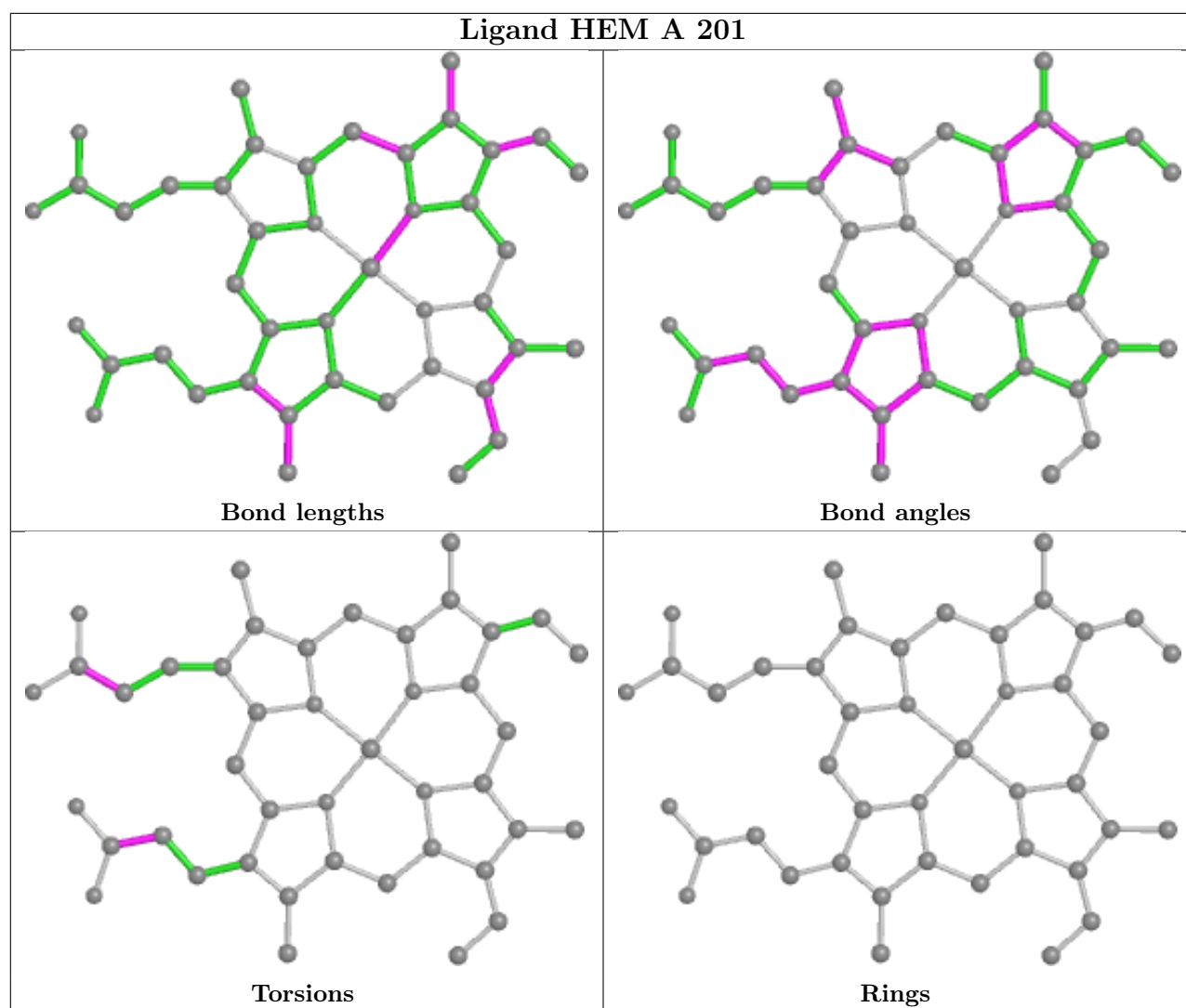
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

## 6 Map visualisation

This section contains visualisations of the EMDB entry EMD-26768. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

### 6.1 Orthogonal projections

This section was not generated.

### 6.2 Central slices

This section was not generated.

### 6.3 Largest variance slices

This section was not generated.

### 6.4 Orthogonal surface views

This section was not generated.

### 6.5 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

## 7 Map analysis ⓘ

This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution ⓘ

This section was not generated.

### 7.2 Volume estimate versus contour level ⓘ

This section was not generated.

### 7.3 Rotationally averaged power spectrum ⓘ

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

## 8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

## 9 Map-model fit

This section was not generated.