



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 24, 2026 – 02:34 PM JST

PDB ID : 9XPD / pdb\_00009xpd  
Title : Crystal Structure of Redesigned HasAsm Variant (51-mutation) with Heme  
Authors : Kim, S.; Sugimoto, H.; Shoji, O.  
Deposited on : 2025-11-16  
Resolution : 1.22 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 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 : 4-5-2 with Phenix2.0  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.48.1

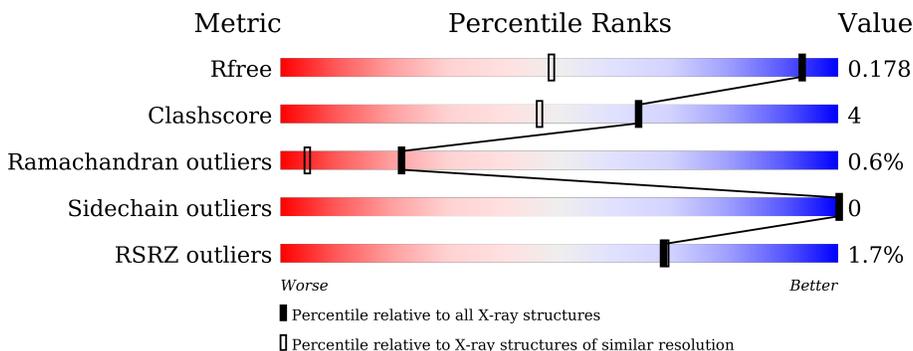
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.22 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))
$R_{free}$	164625	1745 (1.24-1.20)
Clashscore	180529	1895 (1.24-1.20)
Ramachandran outliers	177936	1845 (1.24-1.20)
Sidechain outliers	177891	1844 (1.24-1.20)
RSRZ outliers	164620	1744 (1.24-1.20)

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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	176	

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 2739 atoms, of which 1271 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 Hemophore HasA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	172	2530	819	1223	212	274	2	57	7	0

There are 51 discrepancies between the modelled and reference sequences:

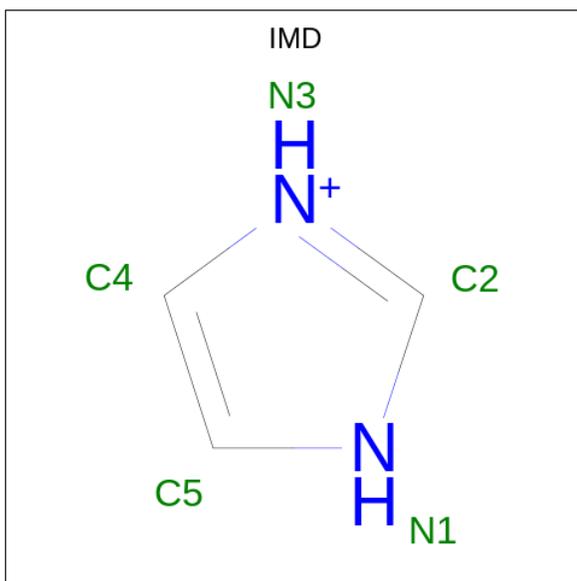
Chain	Residue	Modelled	Actual	Comment	Reference
A	11	LEU	PHE	engineered mutation	UNP Q54450
A	12	ASP	GLY	engineered mutation	UNP Q54450
A	13	ASN	GLY	engineered mutation	UNP Q54450
A	14	LEU	TYR	engineered mutation	UNP Q54450
A	15	THR	SER	engineered mutation	UNP Q54450
A	16	LEU	ILE	engineered mutation	UNP Q54450
A	17	GLN	HIS	engineered mutation	UNP Q54450
A	18	GLU	ASP	engineered mutation	UNP Q54450
A	19	PHE	TYR	engineered mutation	UNP Q54450
A	20	PHE	LEU	engineered mutation	UNP Q54450
A	21	ASN	GLY	engineered mutation	UNP Q54450
A	22	GLU	GLN	engineered mutation	UNP Q54450
A	25	ALA	SER	engineered mutation	UNP Q54450
A	26	SER	THR	engineered mutation	UNP Q54450
A	58	THR	SER	engineered mutation	UNP Q54450
A	61	LYS	ALA	engineered mutation	UNP Q54450
A	62	ASP	ASN	engineered mutation	UNP Q54450
A	63	GLY	GLN	engineered mutation	UNP Q54450
A	88	THR	GLN	engineered mutation	UNP Q54450
A	93	ALA	SER	engineered mutation	UNP Q54450
A	96	ASN	ASP	engineered mutation	UNP Q54450
A	102	SER	ASP	engineered mutation	UNP Q54450
A	103	SER	THR	engineered mutation	UNP Q54450
A	104	THR	SER	engineered mutation	UNP Q54450
A	109	THR	GLN	engineered mutation	UNP Q54450
A	110	GLU	VAL	engineered mutation	UNP Q54450
A	112	GLN	ASP	engineered mutation	UNP Q54450

*Continued on next page...*

Continued from previous page...

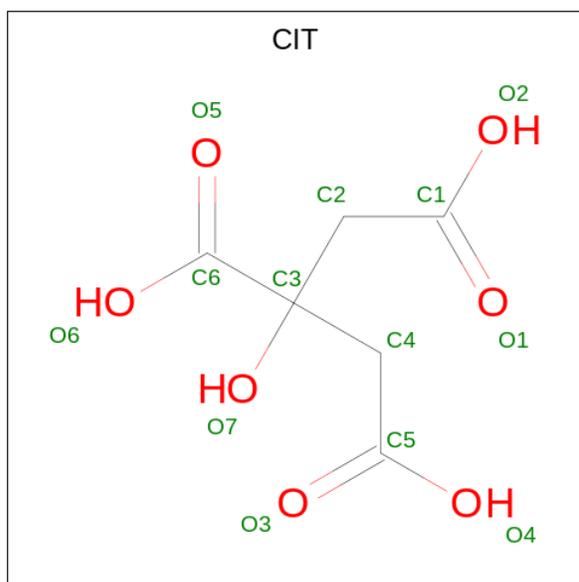
Chain	Residue	Modelled	Actual	Comment	Reference
A	113	PHE	VAL	engineered mutation	UNP Q54450
A	114	ARG	SER	engineered mutation	UNP Q54450
A	115	LEU	PHE	engineered mutation	UNP Q54450
A	119	GLY	ASN	engineered mutation	UNP Q54450
A	120	TRP	LEU	engineered mutation	UNP Q54450
A	121	THR	SER	engineered mutation	UNP Q54450
A	123	GLU	LEU	engineered mutation	UNP Q54450
A	124	ILE	GLN	engineered mutation	UNP Q54450
A	125	SER	ALA	engineered mutation	UNP Q54450
A	126	GLU	GLN	engineered mutation	UNP Q54450
A	144	SER	THR	engineered mutation	UNP Q54450
A	145	THR	GLY	engineered mutation	UNP Q54450
A	146	PRO	ALA	engineered mutation	UNP Q54450
A	148	LEU	GLU	engineered mutation	UNP Q54450
A	149	GLN	THR	engineered mutation	UNP Q54450
A	153	ASN	GLY	engineered mutation	UNP Q54450
A	154	GLN	ILE	engineered mutation	UNP Q54450
A	156	GLN	ASP	engineered mutation	UNP Q54450
A	157	GLN	ASP	engineered mutation	UNP Q54450
A	162	THR	VAL	engineered mutation	UNP Q54450
A	167	MET	ASP	engineered mutation	UNP Q54450
A	168	GLU	GLN	engineered mutation	UNP Q54450
A	169	ILE	VAL	engineered mutation	UNP Q54450
A	170	ARG	ALA	engineered mutation	UNP Q54450

- Molecule 2 is IMIDAZOLE (CCD ID: IMD) (formula: C<sub>3</sub>H<sub>5</sub>N<sub>2</sub>).



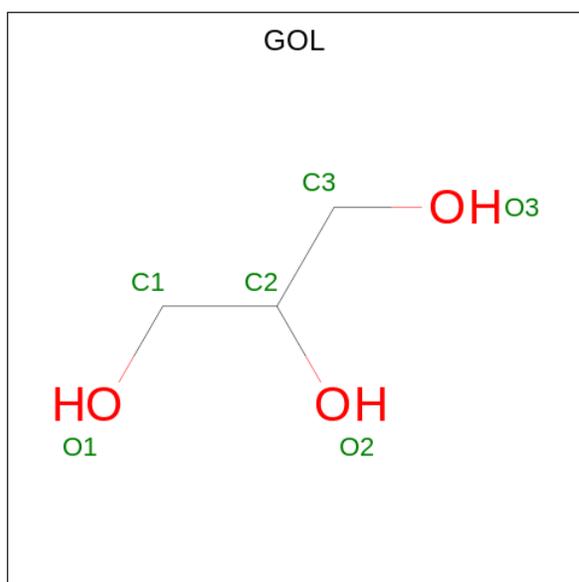
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	N		
2	A	1	10	3	5	2	0	0

- Molecule 3 is CITRIC ACID (CCD ID: CIT) (formula:  $C_6H_8O_7$ ).



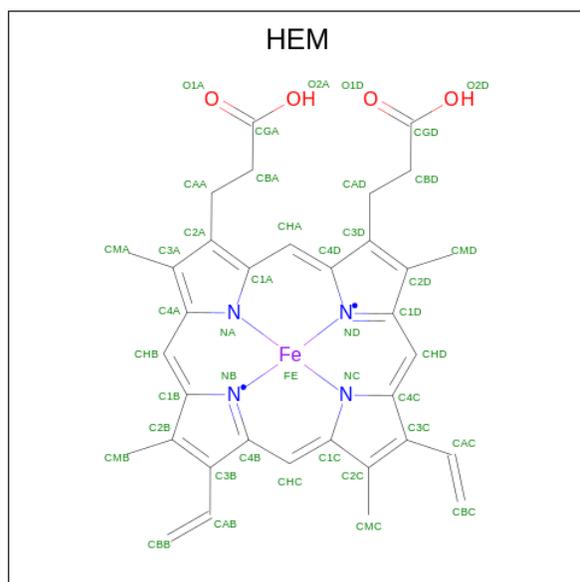
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
3	A	1	18	6	5	7	1	0

- Molecule 4 is GLYCEROL (CCD ID: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	3	0
			14	3	8	3		

- Molecule 5 is PROTOPORPHYRIN IX CONTAINING FE (CCD ID: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
5	A	1	Total	C	Fe	H	N	O	12	0
			73	34	1	30	4	4		

- Molecule 6 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	Na	0	0
			2	2		

- Molecule 7 is water.

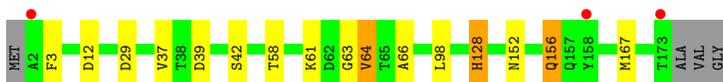
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	92	Total	O	0	0
			92	92		

### 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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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: Hemophore HasA

Chain A:  2% 89% 7% ..



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	30.36Å 140.53Å 80.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.00 – 1.22 35.00 – 1.22	Depositor EDS
% Data completeness (in resolution range)	99.8 (35.00-1.22) 99.8 (35.00-1.22)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 1.22Å)	Xtrriage
Refinement program	REFMAC 5.8.0430 (refmacat 0.4.100)	Depositor
R, $R_{free}$	0.155 , 0.177 0.155 , 0.178	Depositor DCC
$R_{free}$ test set	2633 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.0	Xtrriage
Anisotropy	0.330	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.45 , 31.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	2739	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.15% 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, NA, GOL, IMD, CIT

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	0.72	0/1356	1.07	6/1849 (0.3%)

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	156	GLN	CB-CA-C	-6.73	99.24	110.68
1	A	98	LEU	N-CA-CB	-5.46	102.14	110.65
1	A	3	PHE	N-CA-CB	-5.45	101.36	109.69
1	A	128	HIS	CB-CG-CD2	-5.37	124.23	131.20
1	A	29	ASP	CA-CB-CG	5.20	117.80	112.60
1	A	39	ASP	CA-CB-CG	5.15	117.75	112.60

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	1307	1223	1205	10	0
2	A	5	5	5	0	0
3	A	13	5	5	0	0
4	A	6	8	8	0	0
5	A	43	30	30	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	2	0	0	0	0
7	A	92	0	0	2	0
All	All	1468	1271	1253	10	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 4.

All (10) 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:61:LYS:NZ	7:A:301:HOH:O	2.19	0.73
1:A:12:ASP:HB3	1:A:167:MET:HE3	1.82	0.59
1:A:37:VAL:HG13	1:A:42[B]:SER:OG	2.05	0.56
1:A:128:HIS:HD2	7:A:385:HOH:O	1.87	0.55
1:A:152:ASN:O	1:A:156:GLN:HG3	2.07	0.54
1:A:37:VAL:HG13	1:A:42[B]:SER:HG	1.81	0.46
1:A:37:VAL:CG1	1:A:42[B]:SER:OG	2.63	0.46
1:A:12:ASP:HB3	1:A:167:MET:CE	2.47	0.44
1:A:63:GLY:C	1:A:64[B]:VAL:HG13	2.44	0.43
1:A:58[B]:THR:HG22	1:A:66:ALA:HB2	2.02	0.41

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	177/176 (101%)	172 (97%)	3 (2%)	2 (1%)	<b>12</b>   <b>1</b>

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	64[A]	VAL
1	A	64[B]	VAL

### 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	143/139 (103%)	143 (100%)	0	<a href="#">100</a> <a href="#">100</a>

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 (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	13	ASN
1	A	17	GLN
1	A	128	HIS
1	A	134	GLN
1	A	163	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 oligosaccharides in this entry.

## 5.6 Ligand geometry

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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	CIT	A	202	-	12,12,12	1.06	1 (8%)	17,17,17	1.36	2 (11%)
2	IMD	A	201	-	3,5,5	0.33	0	4,5,5	0.65	0
5	HEM	A	204	1	41,50,50	1.15	4 (9%)	45,82,82	1.70	4 (8%)
4	GOL	A	203	-	5,5,5	0.11	0	5,5,5	0.24	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	CIT	A	202	-	-	6/16/16/16	-
5	HEM	A	204	1	-	1/12/54/54	-
2	IMD	A	201	-	-	-	0/1/1/1
4	GOL	A	203	-	-	4/4/4/4	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	204	HEM	C1A-NA	2.50	1.41	1.36
5	A	204	HEM	C3C-C2C	-2.31	1.37	1.40
5	A	204	HEM	C4D-C3D	-2.29	1.41	1.45
5	A	204	HEM	O2D-CGD	-2.17	1.23	1.30
3	A	202	CIT	O2-C1	-2.06	1.23	1.30

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	204	HEM	O1A-CGA-CBA	-6.18	103.23	123.08
5	A	204	HEM	O2A-CGA-CBA	4.77	129.35	114.03

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	204	HEM	C4C-CHD-C1D	4.53	128.54	122.56
3	A	202	CIT	O5-C6-C3	-3.59	117.17	122.25
5	A	204	HEM	O1D-CGD-CBD	-3.02	113.37	123.08
3	A	202	CIT	O6-C6-C3	2.41	117.23	113.05

There are no chirality outliers.

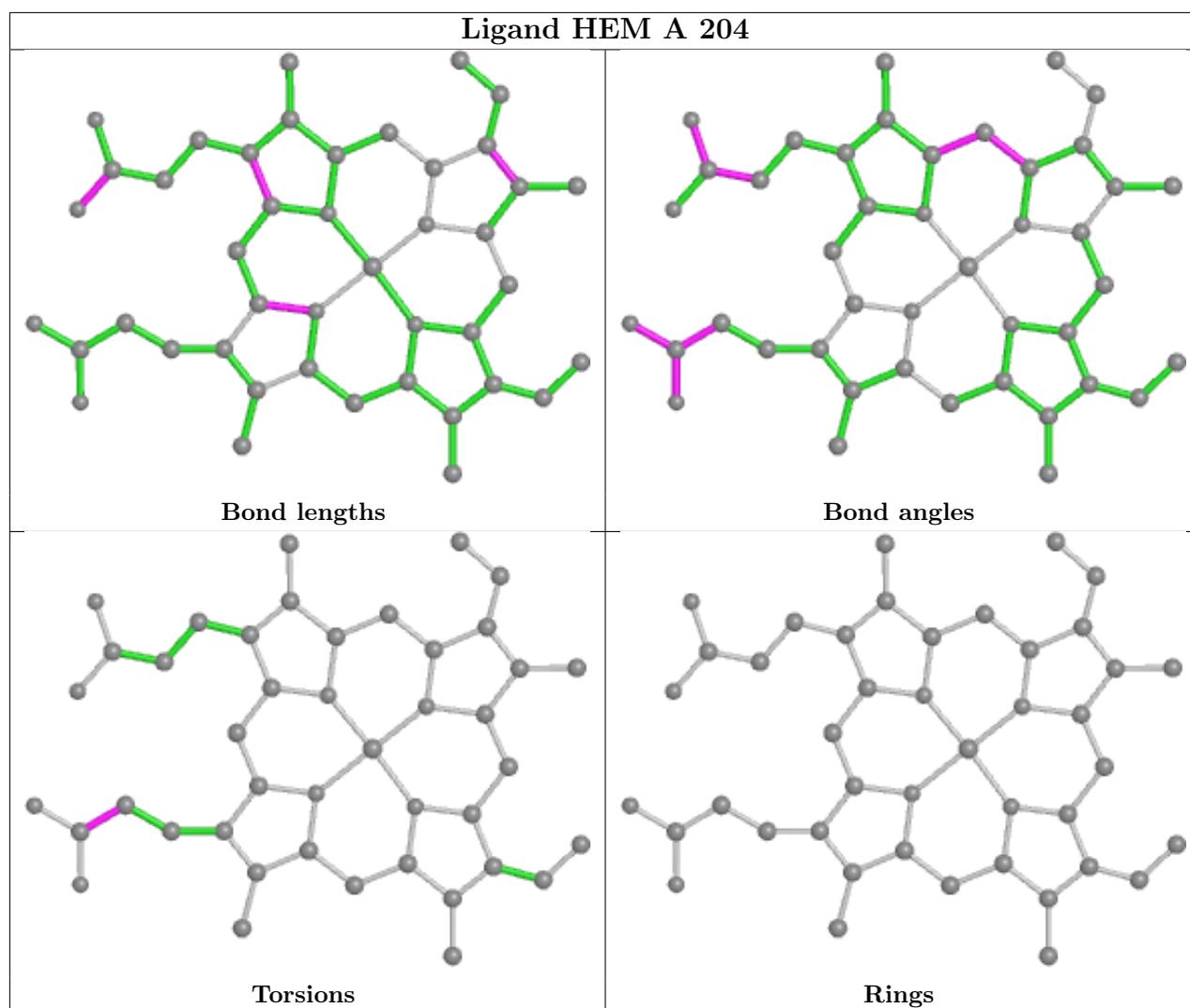
All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	203	GOL	C1-C2-C3-O3
4	A	203	GOL	O2-C2-C3-O3
3	A	202	CIT	C1-C2-C3-C4
4	A	203	GOL	O1-C1-C2-O2
3	A	202	CIT	C4-C3-C6-O6
3	A	202	CIT	C2-C3-C6-O6
5	A	204	HEM	CAA-CBA-CGA-O2A
3	A	202	CIT	C3-C4-C5-O3
3	A	202	CIT	C3-C4-C5-O4
4	A	203	GOL	O1-C1-C2-C3
3	A	202	CIT	C1-C2-C3-C6

There are no ring outliers.

No monomer is involved in short contacts.

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	172/176 (97%)	-0.20	3 (1%) 69 69	5, 16, 29, 40	5 (2%)

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	173	THR	3.1
1	A	158	TYR	2.9
1	A	2	ALA	2.6

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

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

### 6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

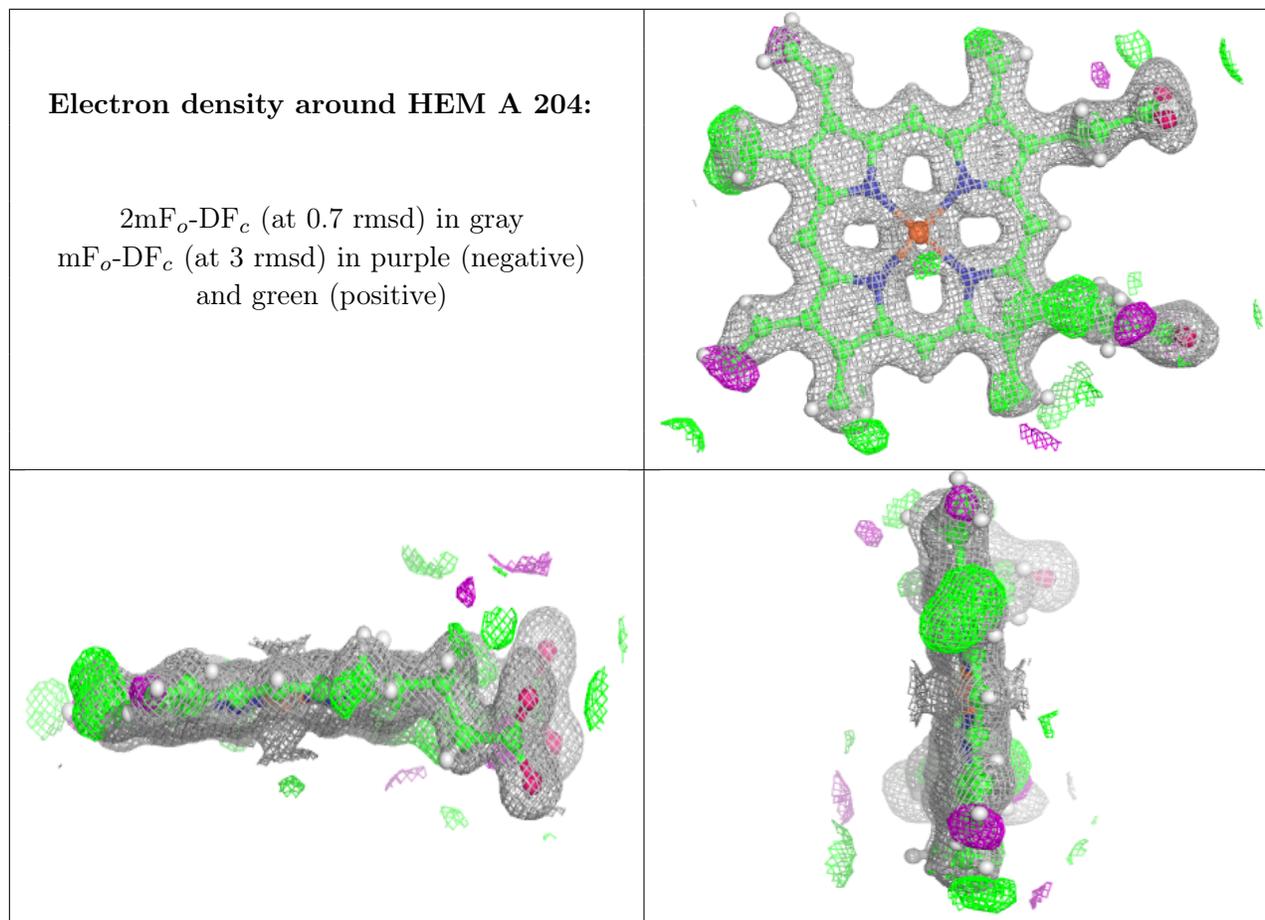
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	GOL	A	203	6/6	0.83	0.14	41,43,45,46	14
6	NA	A	206	1/1	0.93	0.12	37,37,37,37	0
2	IMD	A	201	5/5	0.95	0.09	22,23,23,24	0
5	HEM	A	204	43/43	0.98	0.07	10,14,22,24	12

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CIT	A	202	13/13	0.98	0.07	19,22,26,28	18
6	NA	A	205	1/1	0.99	0.15	22,22,22,22	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.