



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 09:57 am BST

PDB ID : 2WIV  
Title : Cytochrome-P450 XplA heme domain P21  
Authors : Sabbadin, F.; Jackson, R.; Bruce, N.C.; Grogan, G.  
Deposited on : 2009-05-18  
Resolution : 1.90 Å(reported)

This is a wwPDB X-ray Structure 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/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)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

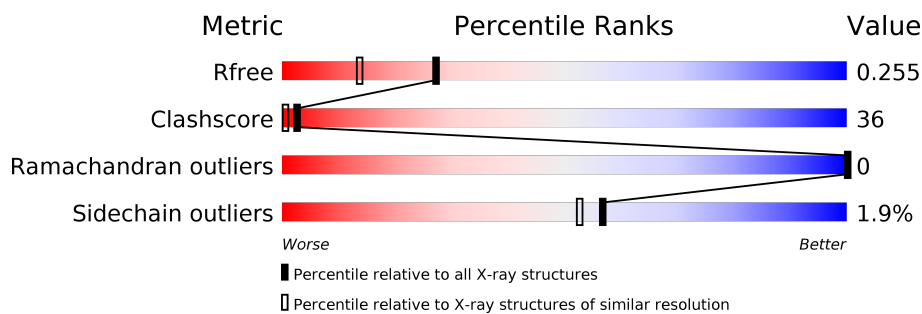
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)

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 on 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	394	
1	B	394	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 6683 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 CYTOCHROME P450-LIKE PROTEIN XPLA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	391	Total	C	N	O	S	30	1	0
			3055	1930	536	577	12			
1	B	391	Total	C	N	O	S	28	1	0
			3056	1929	536	579	12			

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



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

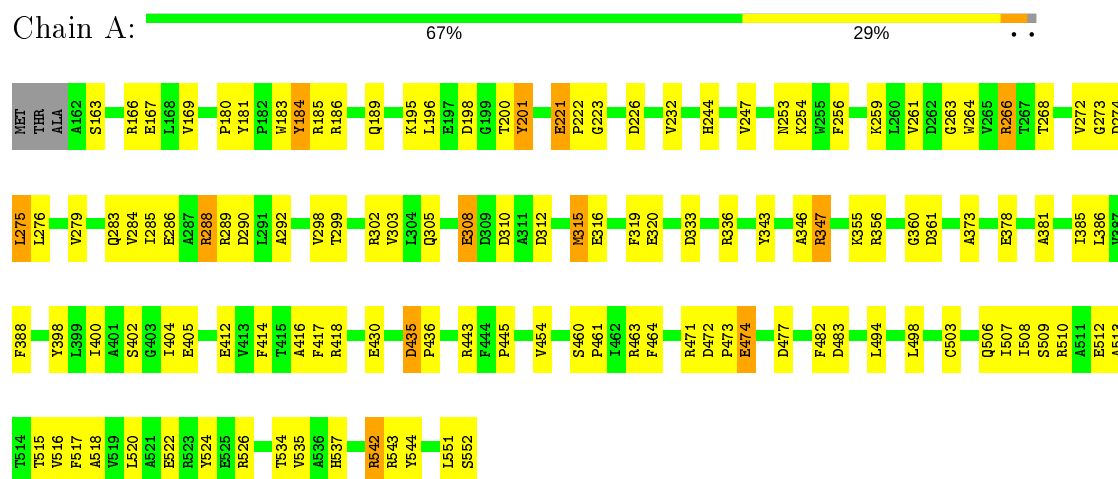
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	253	Total 253	O 253	0	0
3	B	233	Total 233	O 233	0	0

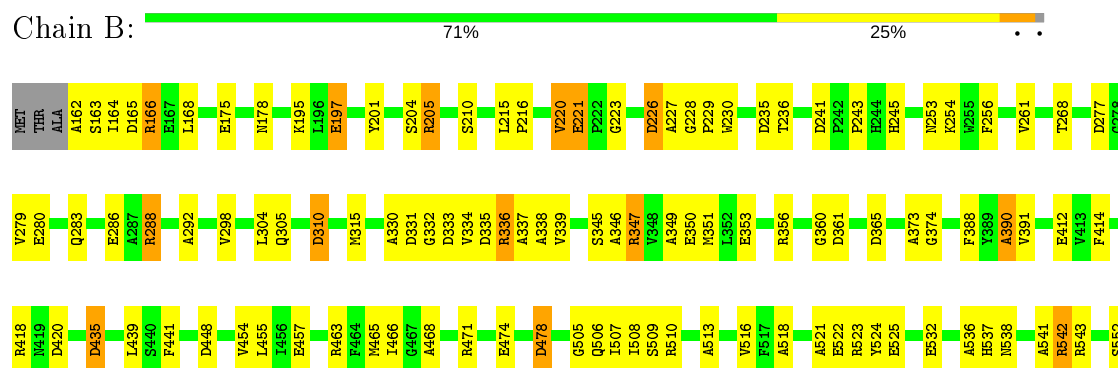
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: CYTOCHROME P450-LIKE PROTEIN XPLA



#### • Molecule 1: CYTOCHROME P450-LIKE PROTEIN XPLA



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.13 Å 123.02 Å 64.61 Å 90.00° 98.54° 90.00°	Depositor
Resolution (Å)	63.89 – 1.90 19.90 – 1.90	Depositor EDS
% Data completeness (in resolution range)	91.5 (63.89-1.90) 91.6 (19.90-1.90)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.89 (at 1.90 Å)	Xtriage
Refinement program	REFMAC 5.5.0082	Depositor
R, $R_{free}$	0.167 , 0.214 0.223 , 0.255	Depositor DCC
$R_{free}$ test set	3104 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.7	Xtriage
Anisotropy	0.278	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 38.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6683	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% 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

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	1.32	11/3132 (0.4%)	1.13	19/4267 (0.4%)
1	B	1.54	10/3133 (0.3%)	1.17	16/4268 (0.4%)
All	All	1.43	21/6265 (0.3%)	1.15	35/8535 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	347	ARG	CD-NE	-55.39	0.52	1.46
1	A	221	GLU	CG-CD	-24.80	1.14	1.51
1	A	474	GLU	CG-CD	-12.17	1.33	1.51
1	B	356	ARG	CB-CG	-11.90	1.20	1.52
1	A	308	GLU	CB-CG	-8.80	1.35	1.52

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	347	ARG	CD-NE-CZ	-35.62	73.73	123.60
1	A	221	GLU	CB-CG-CD	22.97	176.23	114.20
1	B	347	ARG	CG-CD-NE	-13.74	82.94	111.80
1	A	542	ARG	NE-CZ-NH2	-12.76	113.92	120.30
1	A	356	ARG	CA-CB-CG	12.31	140.49	113.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	197	GLU	Sidechain

## 5.2 Too-close contacts ⓘ

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	3055	0	2976	216	2
1	B	3056	0	2971	220	6
2	A	43	0	30	7	0
2	B	43	0	30	2	0
3	A	253	0	0	187	9
3	B	233	0	0	203	1
All	All	6683	0	6007	439	9

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 36.

The worst 5 of 439 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:B:414:PHE:HD2	3:B:2206:HOH:O	1.10	1.34
1:A:247:VAL:HB	3:A:2051:HOH:O	1.20	1.32
1:B:332:GLY:HA3	3:B:2097:HOH:O	1.18	1.32
1:B:478:ASP:HA	3:B:2168:HOH:O	1.17	1.30
1:A:510:ARG:HA	3:A:2222:HOH:O	1.32	1.30

The worst 5 of 9 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:B:243:PRO:CA	3:A:2153:HOH:O[2_645]	0.40	1.80
1:B:243:PRO:C	3:A:2153:HOH:O[2_645]	1.26	0.94
1:B:243:PRO:CB	3:A:2153:HOH:O[2_645]	1.49	0.71

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:186:ARG:CA	3:A:2128:HOH:O[1_455]	1.77	0.43
1:B:243:PRO:N	3:A:2153:HOH:O[2_645]	1.83	0.37

## 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	390/394 (99%)	383 (98%)	7 (2%)	0	100	100
1	B	390/394 (99%)	377 (97%)	13 (3%)	0	100	100
All	All	780/788 (99%)	760 (97%)	20 (3%)	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	319/320 (100%)	312 (98%)	7 (2%)	52	47
1	B	319/320 (100%)	314 (98%)	5 (2%)	62	60
All	All	638/640 (100%)	626 (98%)	12 (2%)	57	53

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	355	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	474	GLU
1	B	310	ASP
1	A	347	ARG
1	B	286	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	178	ASN
1	B	537	HIS
1	B	438	GLN
1	A	506	GLN
1	B	393	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

2 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
2	HEM	B	1553	1	27,50,50	2.28	10 (37%)	17,82,82	1.99	5 (29%)
2	HEM	A	1553	1	27,50,50	2.49	11 (40%)	17,82,82	2.38	8 (47%)

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
2	HEM	B	1553	1	-	0/6/54/54	-
2	HEM	A	1553	1	-	0/6/54/54	-

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1553	HEM	C3C-C2C	-5.45	1.32	1.40
2	B	1553	HEM	C3D-C2D	4.83	1.52	1.37
2	B	1553	HEM	C3C-C2C	-4.77	1.33	1.40
2	A	1553	HEM	C3D-C2D	4.69	1.51	1.37
2	A	1553	HEM	C3B-C2B	-3.76	1.35	1.40

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1553	HEM	CBD-CAD-C3D	-5.21	102.88	112.48
2	B	1553	HEM	CBD-CAD-C3D	-4.48	104.22	112.48
2	A	1553	HEM	CMA-C3A-C4A	-3.98	122.34	128.46
2	A	1553	HEM	CMA-C3A-C2A	3.41	131.37	124.94
2	B	1553	HEM	CMA-C3A-C4A	-3.31	123.37	128.46

There are no chirality outliers.

There are no torsion outliers.

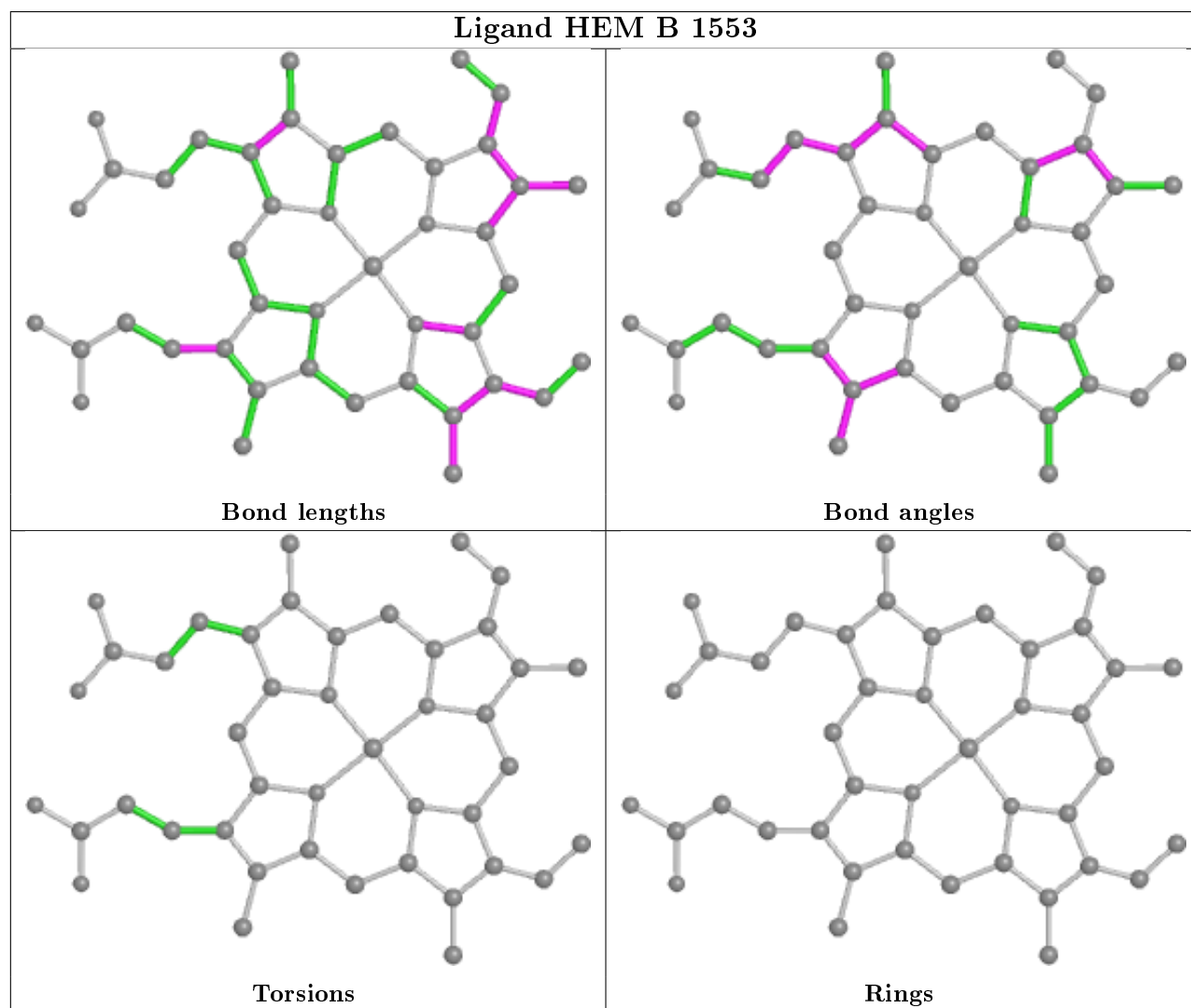
There are no ring outliers.

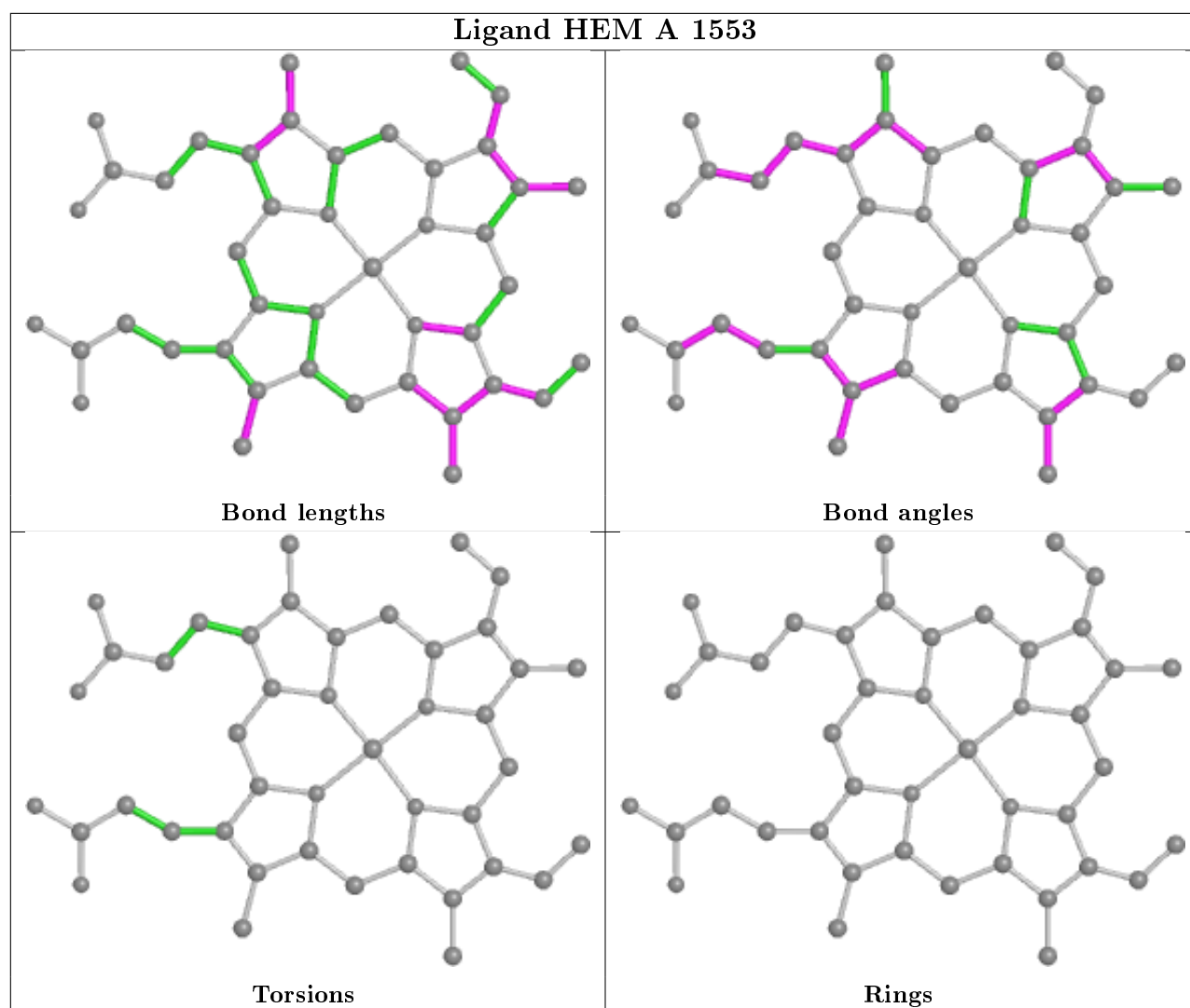
2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1553	HEM	2	0
2	A	1553	HEM	7	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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

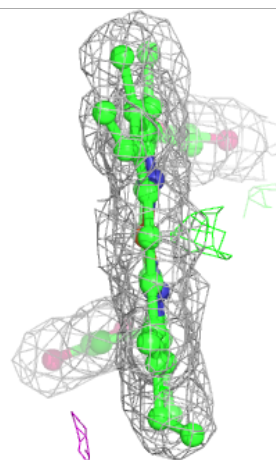
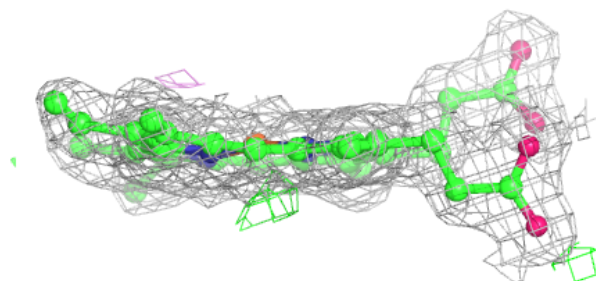
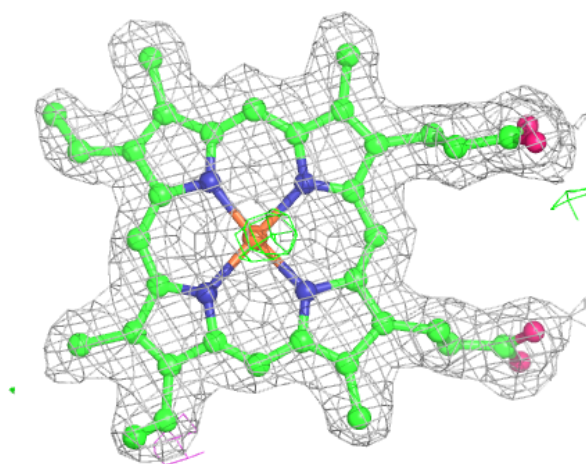
### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

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.

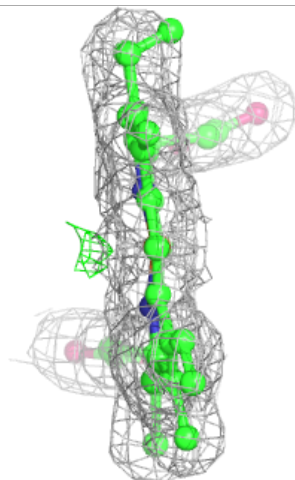
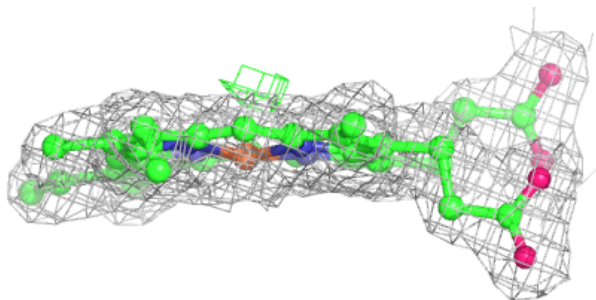
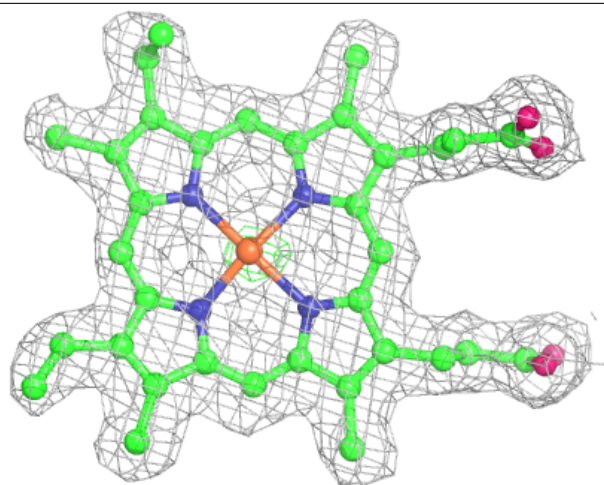
**Electron density around HEM B 1553:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around HEM A 1553:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.