



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 22, 2024 – 06:04 PM EDT

PDB ID : 6IQ5  
Title : Crystal Structure of CYP1B1 and Inhibitor Having Azide Group  
Authors : Kubo, M.; Yamamoto, K.; Itoh, T.  
Deposited on : 2018-11-06  
Resolution : 3.70 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 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	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
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)	:	2.37.1

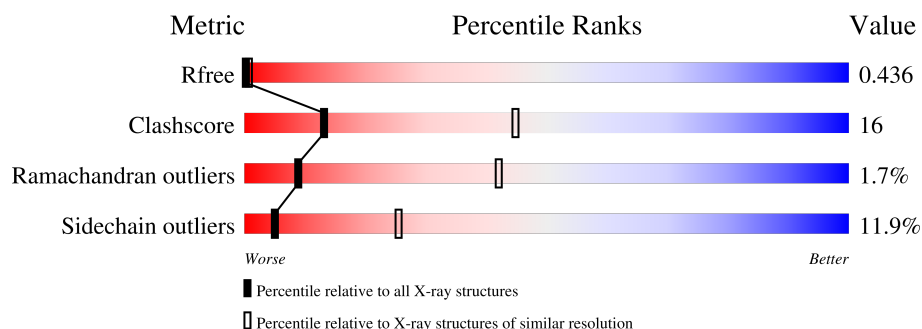
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.70 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	1049 (3.88-3.52)
Clashscore	141614	1027 (3.86-3.54)
Ramachandran outliers	138981	1069 (3.88-3.52)
Sidechain outliers	138945	1065 (3.88-3.52)

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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	463	
1	B	463	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7186 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.

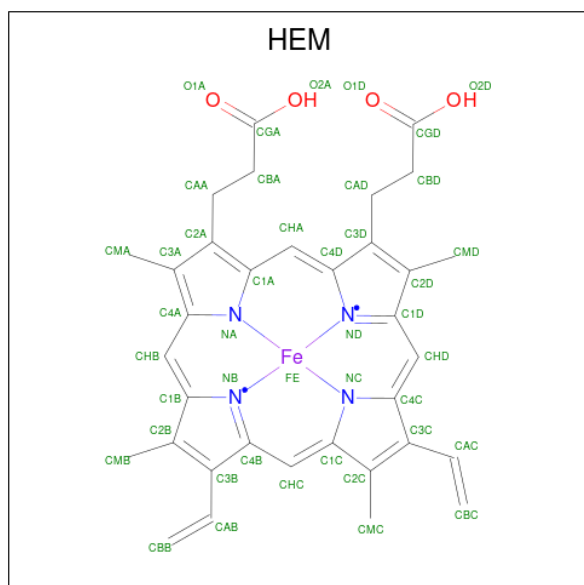
- Molecule 1 is a protein called Cytochrome P450 1B1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	459	Total	C	N	O	S	0	0	0
			3650	2322	656	652	20			
1	B	437	Total	C	N	O	S	0	0	0
			3423	2187	612	605	19			

There are 4 discrepancies between the modelled and reference sequences:

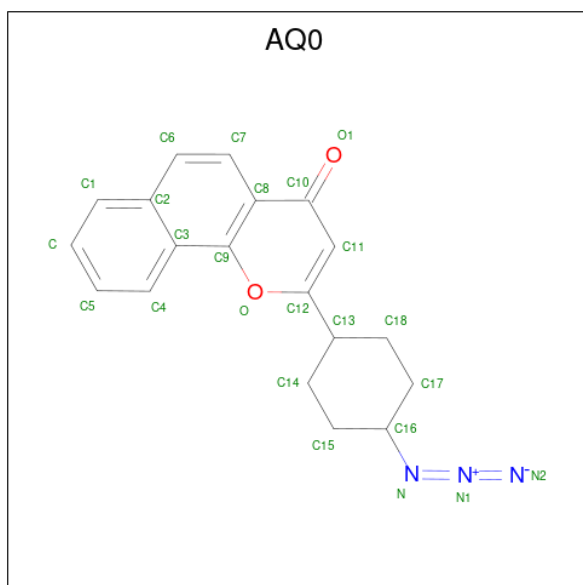
Chain	Residue	Modelled	Actual	Comment	Reference
A	119	SER	ALA	variant	UNP Q16678
A	432	VAL	LEU	variant	UNP Q16678
B	119	SER	ALA	variant	UNP Q16678
B	432	VAL	LEU	variant	UNP Q16678

- 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 43	C 34	Fe 1	N 4	O 4	0	0
2	B	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

- Molecule 3 is 2-(cis-4-azidocyclohexyl)-4H-naphtho[1,2-b]pyran-4-one (three-letter code: AQ0) (formula: C<sub>19</sub>H<sub>17</sub>N<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			24	19	3	2		

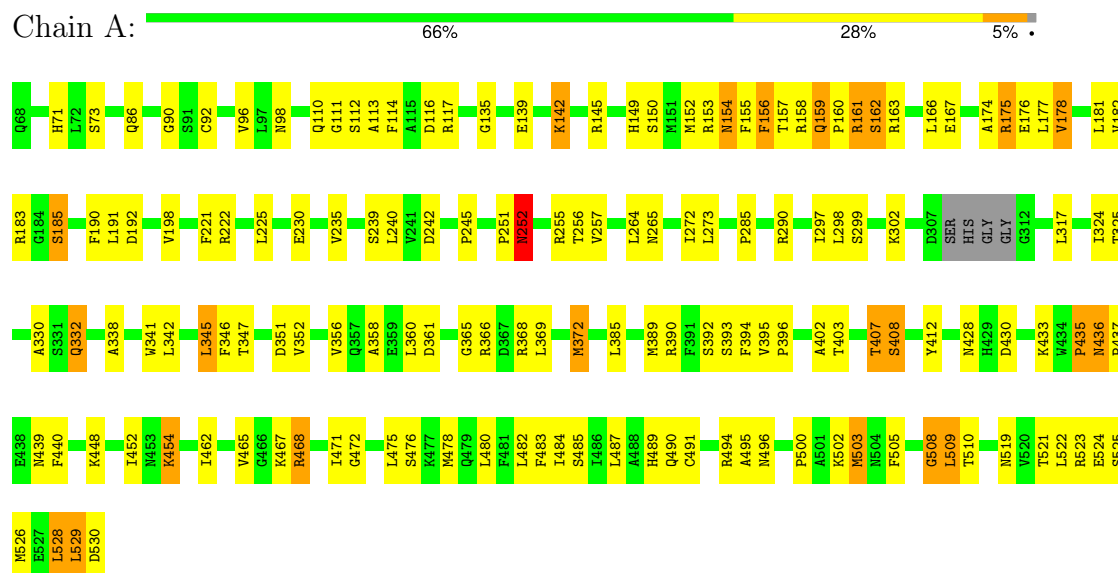
- Molecule 4 is water.

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

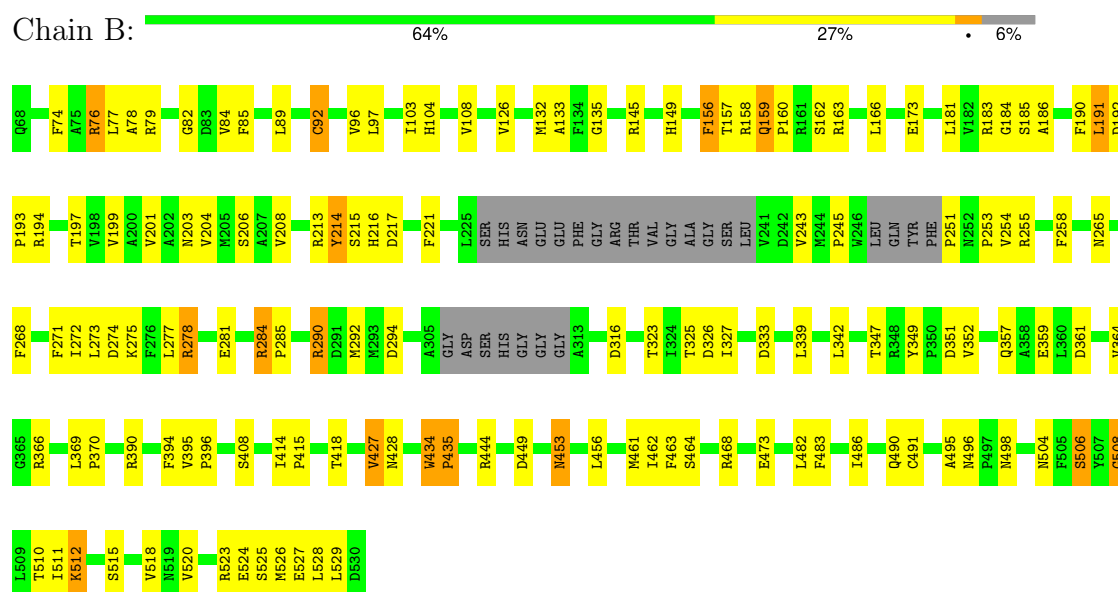
### 3 Residue-property plots

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: Cytochrome P450 1B1



#### • Molecule 1: Cytochrome P450 1B1



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	135.19Å 135.19Å 302.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.80 – 3.70 47.80 – 3.70	Depositor EDS
% Data completeness (in resolution range)	98.8 (47.80-3.70) 99.4 (47.80-3.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	7.06 (at 3.67Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, $R_{free}$	0.356 , 0.396 0.419 , 0.436	Depositor DCC
$R_{free}$ test set	1472 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	111.3	Xtriage
Anisotropy	0.012	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 310.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.025 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.75	EDS
Total number of atoms	7186	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	316.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, AQ0

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	0.59	0/3739	0.83	0/5066
1	B	0.62	0/3506	0.86	0/4759
All	All	0.61	0/7245	0.84	0/9825

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	3650	0	3606	129	0
1	B	3423	0	3349	102	0
2	A	43	0	30	4	0
2	B	43	0	30	5	0
3	A	24	0	0	1	0
4	A	2	0	0	0	0
4	B	1	0	0	0	0
All	All	7186	0	7015	231	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 16.

The worst 5 of 231 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:526:MET:CE	1:A:529:LEU:HD21	1.69	1.22
1:A:526:MET:HE1	1:A:529:LEU:HD21	1.22	1.10
1:B:327:ILE:HG23	2:B:900:HEM:HBC1	1.17	1.10
1:A:526:MET:CE	1:A:529:LEU:CD2	2.38	1.00
1:A:489:HIS:HA	1:A:523:ARG:NH1	1.77	0.99

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	455/463 (98%)	393 (86%)	54 (12%)	8 (2%)	8	41
1	B	429/463 (93%)	367 (86%)	55 (13%)	7 (2%)	9	43
All	All	884/926 (96%)	760 (86%)	109 (12%)	15 (2%)	9	42

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	285	PRO
1	A	436	ASN
1	A	500	PRO
1	B	284	ARG
1	B	435	PRO

### 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	394/397 (99%)	336 (85%)	58 (15%)	3	18
1	B	362/397 (91%)	330 (91%)	32 (9%)	10	38
All	All	756/794 (95%)	666 (88%)	90 (12%)	5	25

5 of 90 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	528	LEU
1	B	278	ARG
1	B	92	CYS
1	B	173	GLU
1	B	342	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	71	HIS
1	B	98	ASN
1	B	423	ASN
1	B	265	ASN
1	A	357	GLN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry

3 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	A	901	1	42,50,50	1.40	7 (16%)	46,82,82	1.80	14 (30%)
3	AQ0	A	902	-	27,27,27	1.49	4 (14%)	31,38,38	2.31	7 (22%)
2	HEM	B	900	1	42,50,50	1.46	9 (21%)	46,82,82	1.67	12 (26%)

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	A	901	1	-	0/12/54/54	-
3	AQ0	A	902	-	-	4/7/17/17	1/4/4/4
2	HEM	B	900	1	-	0/12/54/54	-

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	902	AQ0	C8-C9	4.19	1.48	1.39
2	B	900	HEM	C1B-NB	-3.76	1.33	1.40
2	A	901	HEM	C1B-NB	-3.70	1.33	1.40
2	A	901	HEM	C4D-ND	-3.21	1.34	1.40
3	A	902	AQ0	C3-C2	3.15	1.48	1.43

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	902	AQ0	C16-N-N1	8.39	133.12	115.10
3	A	902	AQ0	O-C9-C3	5.03	122.79	114.77
3	A	902	AQ0	C9-C8-C10	-4.46	115.84	119.55
2	A	901	HEM	CHA-C4D-ND	4.11	129.47	124.37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	901	HEM	CHC-C4B-NB	4.06	128.80	124.44

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	902	AQ0	O-C12-C13-C14
3	A	902	AQ0	C11-C12-C13-C14
3	A	902	AQ0	C15-C16-N-N1
3	A	902	AQ0	O-C12-C13-C18

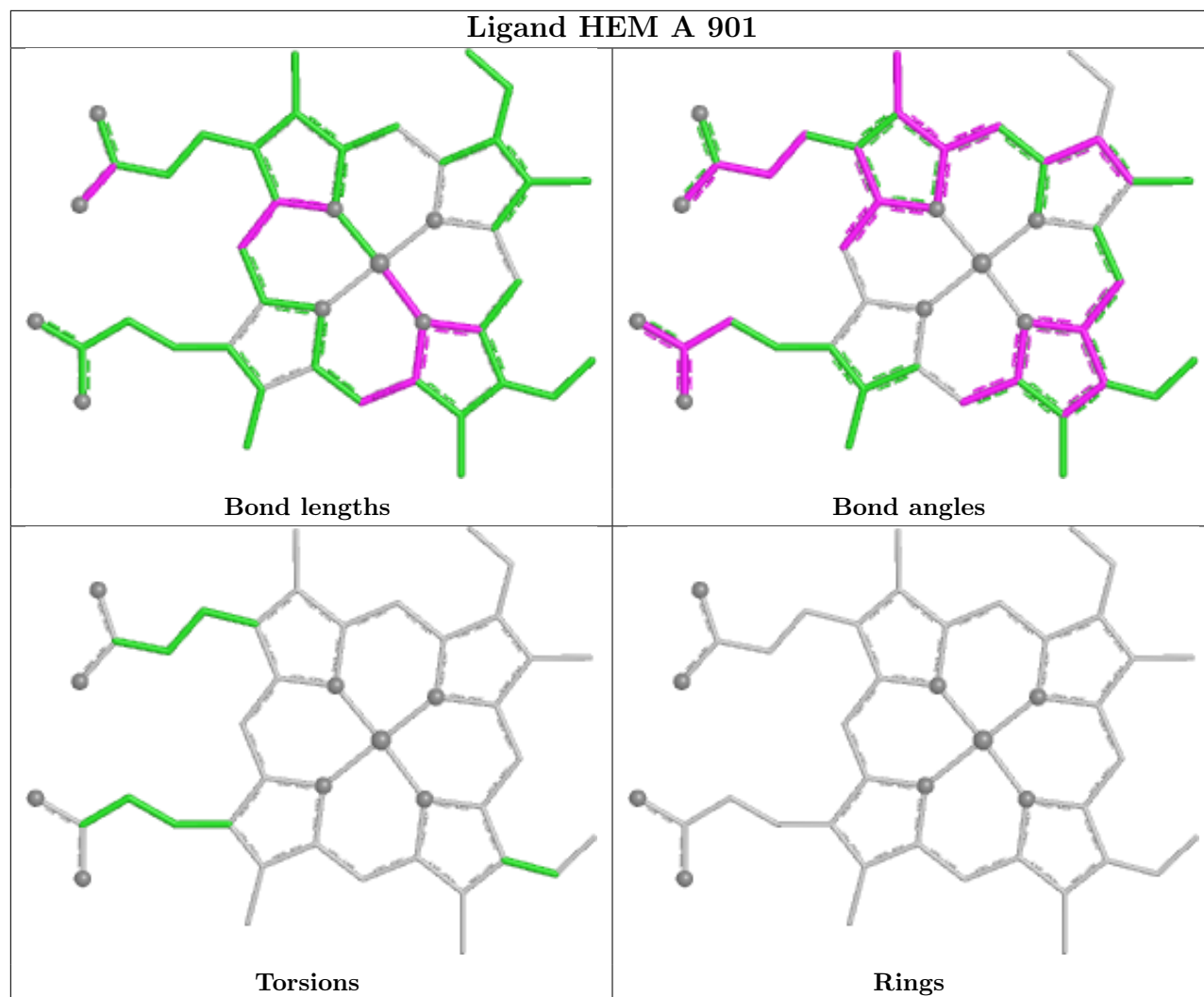
All (1) ring outliers are listed below:

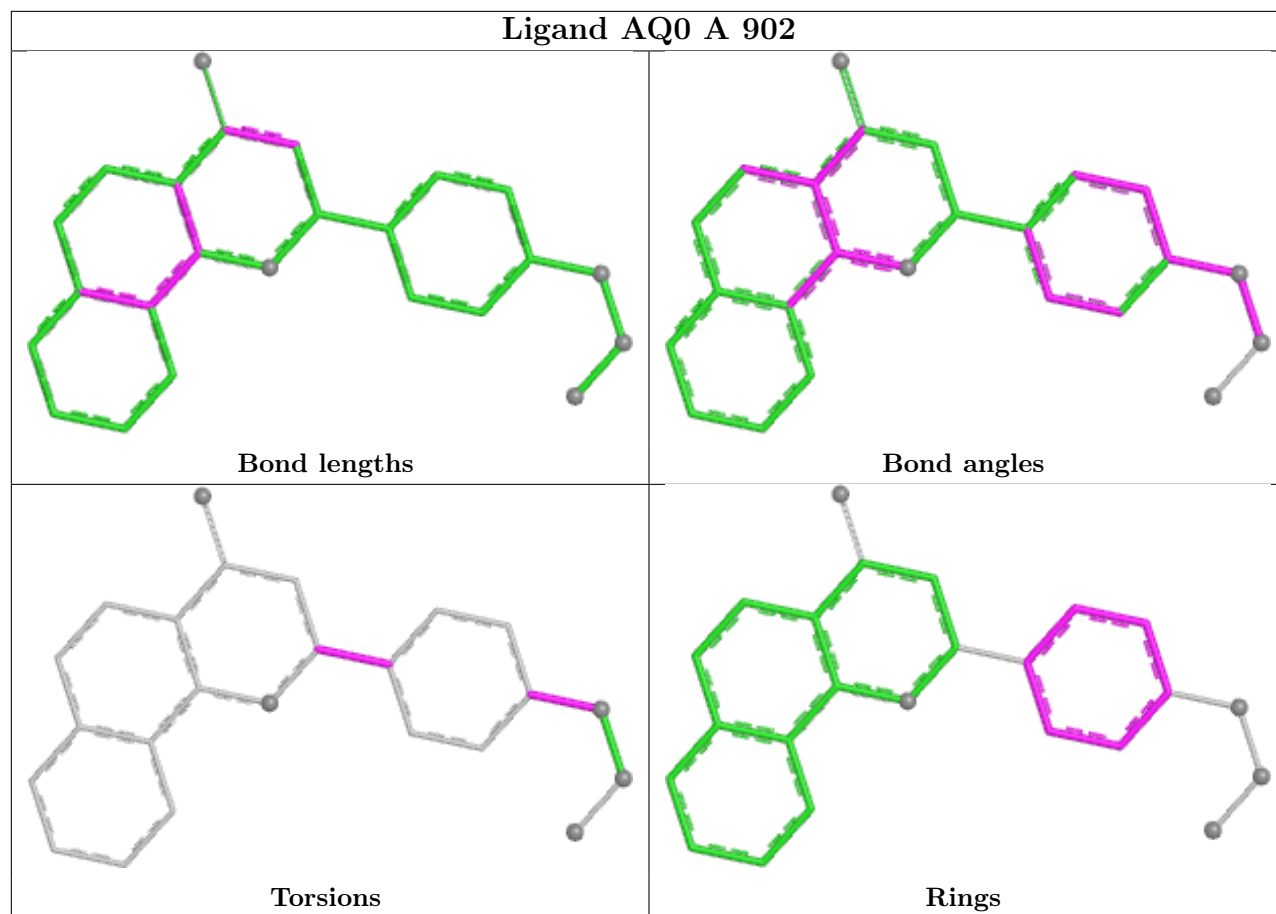
Mol	Chain	Res	Type	Atoms
3	A	902	AQ0	C13-C14-C15-C16-C17-C18

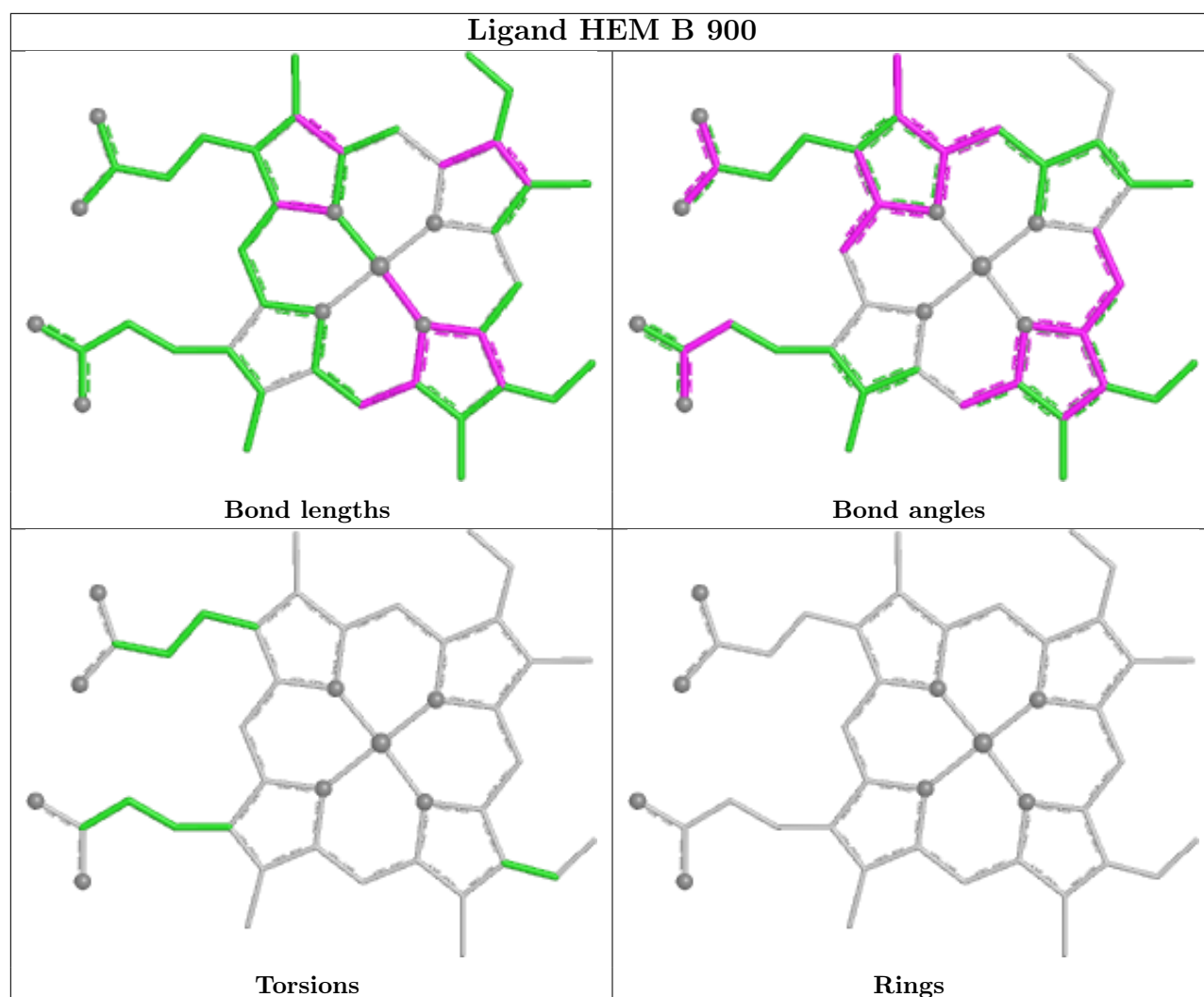
3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	901	HEM	4	0
3	A	902	AQ0	1	0
2	B	900	HEM	5	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.

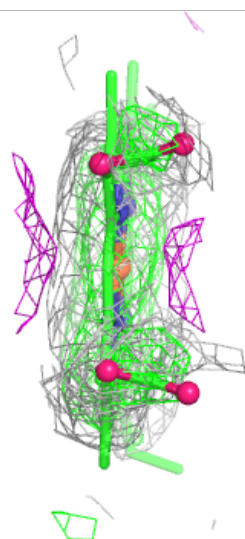
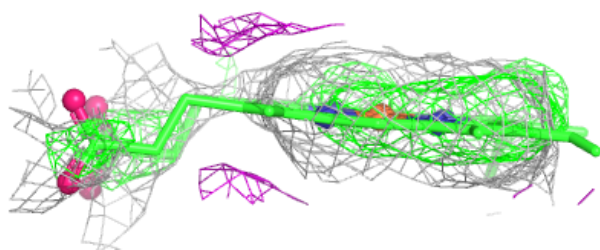
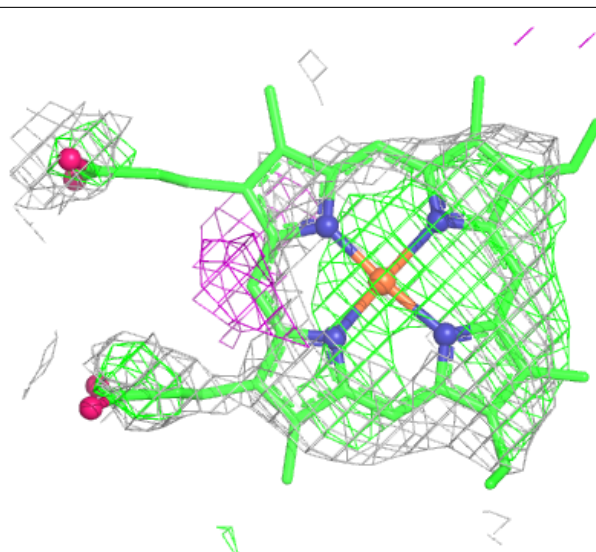
### 6.4 Ligands

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 A 901:**

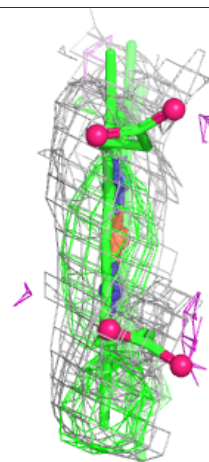
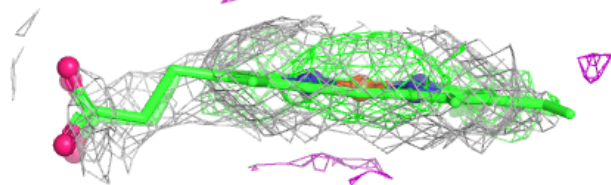
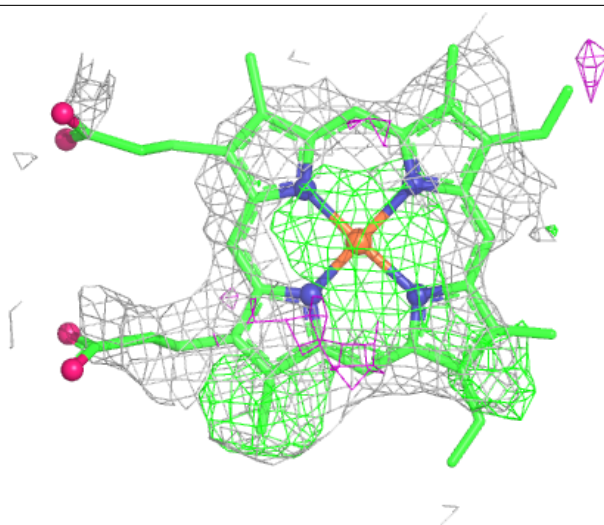
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





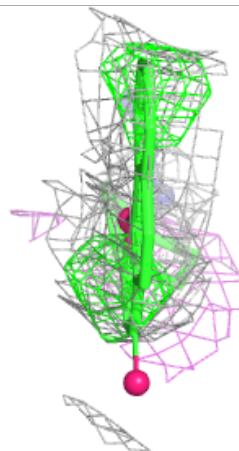
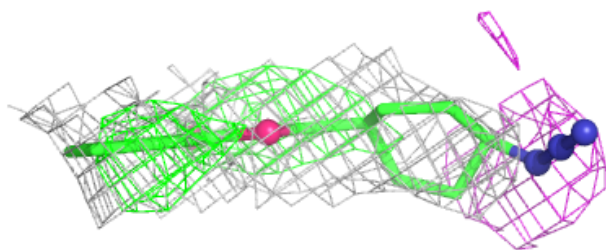
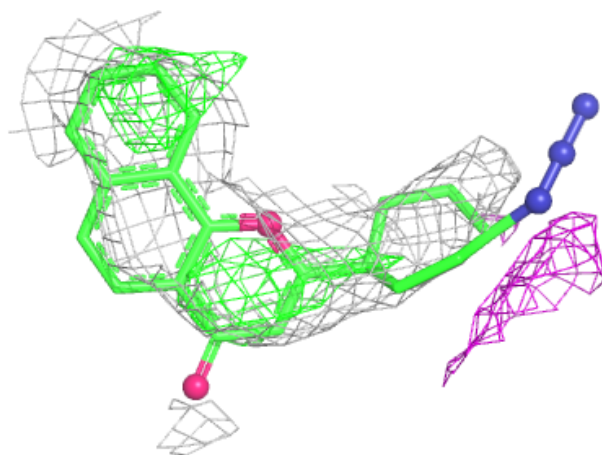
**Electron density around HEM B 900:**

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



**Electron density around AQ0 A 902:**

$2mF_o-DF_c$  (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.