



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 25, 2024 – 01:34 am BST

PDB ID : 9G9S  
Title : Crystal structure of PbdA bound to veratrate  
Authors : Hinchen, D.J.; Wolf, M.E.; Eltis, L.D.; McGeehan, J.E.  
Deposited on : 2024-07-25  
Resolution : 1.85 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 : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.002 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.38.2

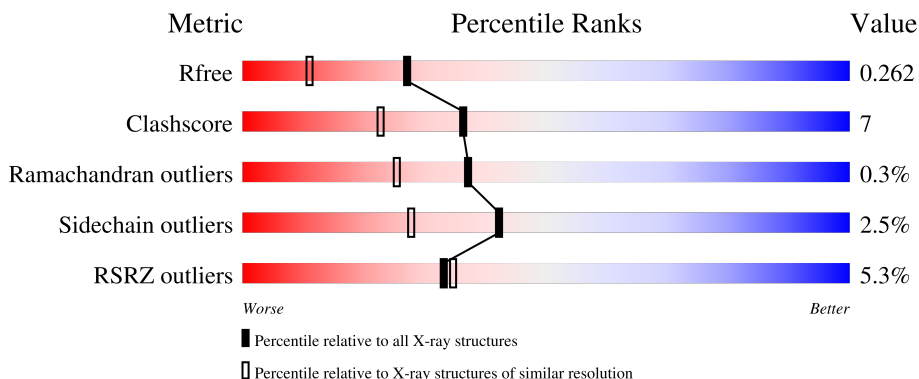
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.85 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	3097 (1.86-1.86)
Clashscore	180529	3359 (1.86-1.86)
Ramachandran outliers	177936	3335 (1.86-1.86)
Sidechain outliers	177891	3335 (1.86-1.86)
RSRZ outliers	164620	3097 (1.86-1.86)

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	420	 7% 79% 14% • 6%
1	B	420	 3% 81% 11% • 6%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 12744 atoms, of which 6052 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 CYP199.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	393	6001	1924	2966	550	557	4	73	2	0
1	B	393	6037	1935	2992	551	555	4	76	4	0

There are 40 discrepancies between the modelled and reference sequences:

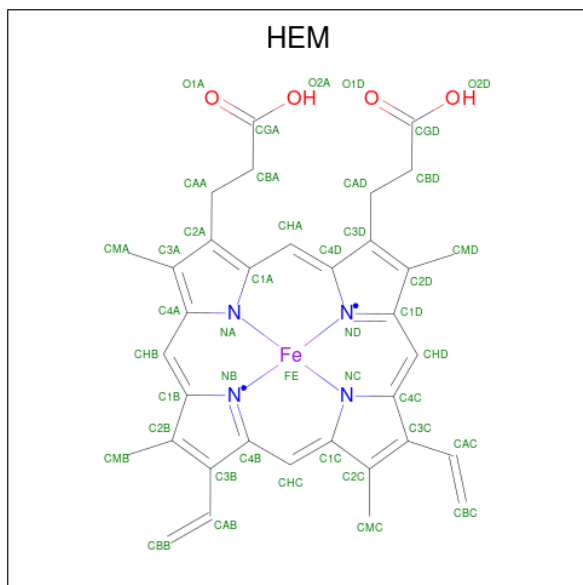
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP Q0SCI3
A	-18	GLY	-	expression tag	UNP Q0SCI3
A	-17	SER	-	expression tag	UNP Q0SCI3
A	-16	SER	-	expression tag	UNP Q0SCI3
A	-15	HIS	-	expression tag	UNP Q0SCI3
A	-14	HIS	-	expression tag	UNP Q0SCI3
A	-13	HIS	-	expression tag	UNP Q0SCI3
A	-12	HIS	-	expression tag	UNP Q0SCI3
A	-11	HIS	-	expression tag	UNP Q0SCI3
A	-10	SER	-	expression tag	UNP Q0SCI3
A	-9	SER	-	expression tag	UNP Q0SCI3
A	-8	GLY	-	expression tag	UNP Q0SCI3
A	-7	GLU	-	expression tag	UNP Q0SCI3
A	-6	ASN	-	expression tag	UNP Q0SCI3
A	-5	LEU	-	expression tag	UNP Q0SCI3
A	-4	TYR	-	expression tag	UNP Q0SCI3
A	-3	PHE	-	expression tag	UNP Q0SCI3
A	-2	GLN	-	expression tag	UNP Q0SCI3
A	-1	GLY	-	expression tag	UNP Q0SCI3
A	0	HIS	-	expression tag	UNP Q0SCI3
B	-19	MET	-	initiating methionine	UNP Q0SCI3
B	-18	GLY	-	expression tag	UNP Q0SCI3
B	-17	SER	-	expression tag	UNP Q0SCI3
B	-16	SER	-	expression tag	UNP Q0SCI3
B	-15	HIS	-	expression tag	UNP Q0SCI3

*Continued on next page...*

Continued from previous page...

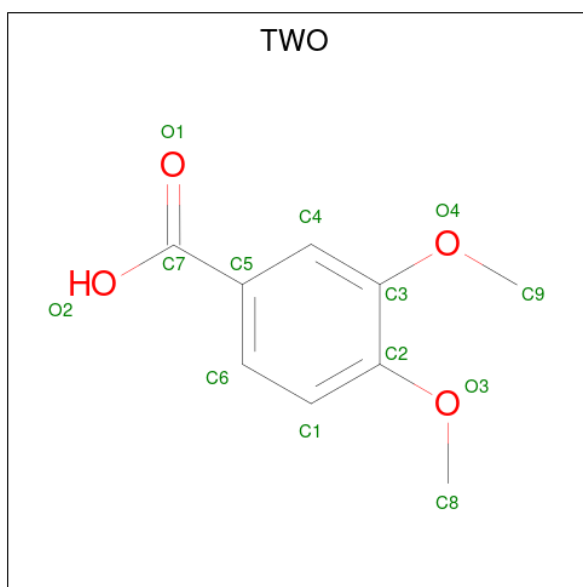
Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	HIS	-	expression tag	UNP Q0SCI3
B	-13	HIS	-	expression tag	UNP Q0SCI3
B	-12	HIS	-	expression tag	UNP Q0SCI3
B	-11	HIS	-	expression tag	UNP Q0SCI3
B	-10	SER	-	expression tag	UNP Q0SCI3
B	-9	SER	-	expression tag	UNP Q0SCI3
B	-8	GLY	-	expression tag	UNP Q0SCI3
B	-7	GLU	-	expression tag	UNP Q0SCI3
B	-6	ASN	-	expression tag	UNP Q0SCI3
B	-5	LEU	-	expression tag	UNP Q0SCI3
B	-4	TYR	-	expression tag	UNP Q0SCI3
B	-3	PHE	-	expression tag	UNP Q0SCI3
B	-2	GLN	-	expression tag	UNP Q0SCI3
B	-1	GLY	-	expression tag	UNP Q0SCI3
B	0	HIS	-	expression tag	UNP Q0SCI3

- 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	
			Total	C	Fe	H	N			O
2	A	1	73	34	1	30	4	4	0	0
2	B	1	73	34	1	30	4	4	0	0

- Molecule 3 is 3,4-dimethoxybenzoic acid (three-letter code: TWO) (formula:  $C_9H_{10}O_4$ ) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
3	A	1	22	9	9	4	0	0
3	B	1	22	9	9	4	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

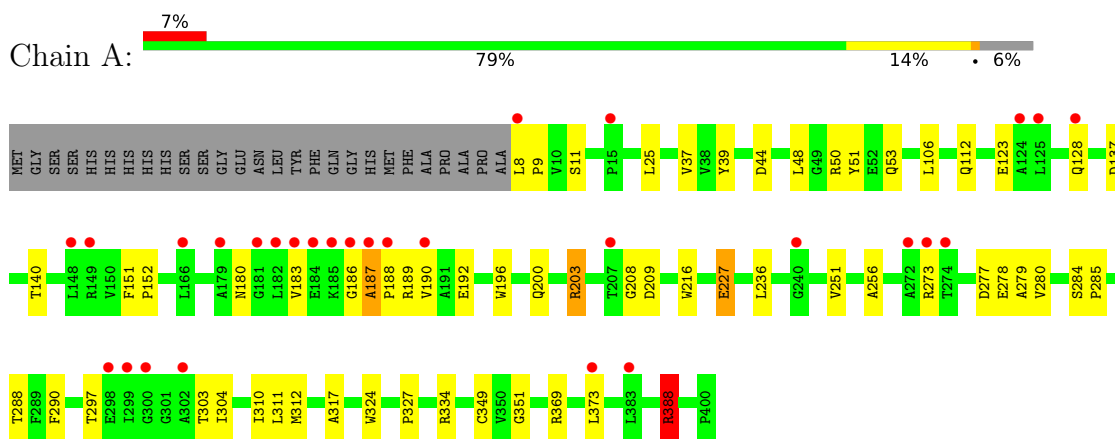
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	229	Total 229	O 229	0	0
5	B	259	Total 259	O 259	0	0

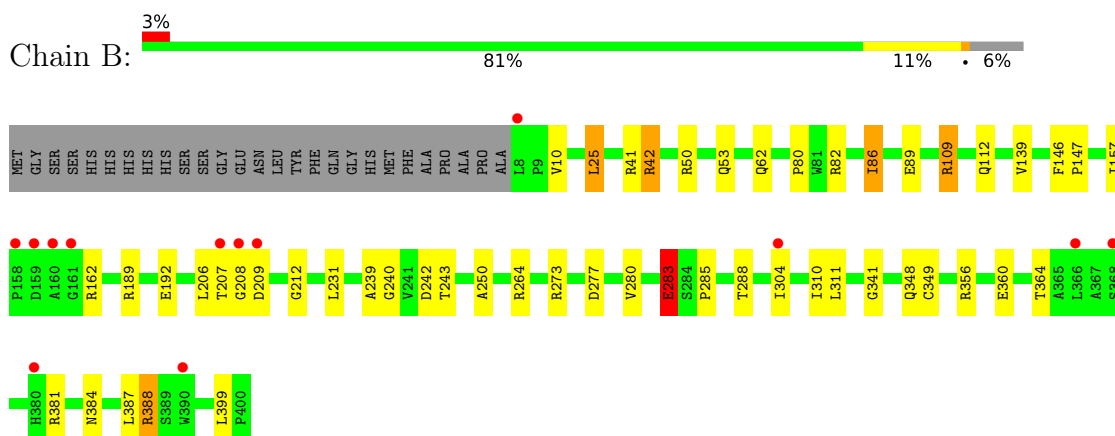
### 3 Residue-property plots [i](#)

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: Cytochrome P450 CYP199



- Molecule 1: Cytochrome P450 CYP199



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.00Å 55.36Å 64.69Å 98.68° 95.00° 99.01°	Depositor
Resolution (Å)	63.54 – 1.85 63.54 – 1.85	Depositor EDS
% Data completeness (in resolution range)	97.5 (63.54-1.85) 95.0 (63.54-1.85)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.24 (at 1.86Å)	Xtrriage
Refinement program	REFMAC 5.8.0430, REFMAC 5.8.0430	Depositor
R, $R_{free}$	0.208 , 0.258 0.211 , 0.262	Depositor DCC
$R_{free}$ test set	2837 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.0	Xtrriage
Anisotropy	0.224	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.46 , 44.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	12744	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.22% 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: TWO, HEM, GOL

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	0.45	0/3125	0.91	6/4274 (0.1%)
1	B	0.46	0/3146	0.88	4/4302 (0.1%)
All	All	0.45	0/6271	0.90	10/8576 (0.1%)

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	A	0	1
1	B	0	3
All	All	0	4

There are no bond length outliers.

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	203	ARG	NE-CZ-NH1	9.76	125.18	120.30
1	A	203	ARG	NE-CZ-NH2	-8.21	116.19	120.30
1	A	388	ARG	NE-CZ-NH2	-7.54	116.53	120.30
1	A	203	ARG	CD-NE-CZ	6.14	132.19	123.60
1	B	25	LEU	CB-CG-CD1	5.52	120.39	111.00
1	A	227	GLU	CG-CD-OE2	-5.51	107.29	118.30
1	B	283	GLU	CG-CD-OE2	-5.47	107.35	118.30
1	B	388	ARG	NE-CZ-NH2	-5.33	117.63	120.30
1	B	89	GLU	CB-CA-C	5.18	120.77	110.40
1	A	227	GLU	CG-CD-OE1	5.14	128.58	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	388	ARG	Sidechain
1	B	109	ARG	Sidechain
1	B	381[A]	ARG	Sidechain
1	B	381[B]	ARG	Sidechain

## 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	3035	2966	2933	43	0
1	B	3045	2992	2945	38	0
2	A	43	30	30	6	0
2	B	43	30	30	6	0
3	A	13	9	9	0	0
3	B	13	9	9	1	0
4	B	12	16	16	2	0
5	A	229	0	0	10	2
5	B	259	0	0	10	2
All	All	6692	6052	5972	87	2

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 7.

All (87) 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:250:ALA:HB1	1:B:283:GLU:HG3	1.45	0.97
1:A:273:ARG:NH2	5:A:601:HOH:O	2.00	0.95
1:A:208:GLY:O	5:A:602:HOH:O	2.03	0.76
1:A:209:ASP:OD1	5:A:603:HOH:O	2.04	0.74
1:A:44:ASP:OD2	5:A:604:HOH:O	2.05	0.74
1:B:208:GLY:O	5:B:601:HOH:O	2.07	0.71
1:B:42:ARG:CG	1:B:42:ARG:HH11	2.04	0.70
2:A:501:HEM:HMC2	2:A:501:HEM:HBC2	1.75	0.69
1:B:42:ARG:HH11	1:B:42:ARG:HG3	1.57	0.69

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:285:PRO:HD2	2:B:501:HEM:HBC1	1.73	0.69
1:B:360:GLU:O	1:B:364:THR:HG23	1.96	0.65
1:B:250:ALA:CB	1:B:283:GLU:HG3	2.23	0.65
1:A:137:ASP:HB3	1:A:140:THR:OG1	1.98	0.63
1:B:304:ILE:HD13	1:B:310:ILE:CG2	2.29	0.63
1:A:304:ILE:HD13	1:A:310:ILE:CG2	2.30	0.61
1:A:196:TRP:O	1:A:200:GLN:HG2	2.01	0.60
1:B:273:ARG:NH2	5:B:606:HOH:O	2.31	0.60
1:B:86[A]:ILE:CD1	1:B:231:LEU:HD13	2.33	0.59
2:A:501:HEM:HBB2	2:A:501:HEM:HMB2	1.86	0.58
1:A:187:ALA:HB3	1:A:188:PRO:HD3	1.86	0.58
2:A:501:HEM:HBC2	2:A:501:HEM:CMC	2.34	0.58
1:A:37:VAL:HG22	1:A:48:LEU:CD2	2.36	0.55
1:B:62:GLN:NE2	5:B:602:HOH:O	2.22	0.55
1:A:25:LEU:HD21	5:A:750:HOH:O	2.07	0.55
1:B:146:PHE:HB3	1:B:147:PRO:HD3	1.90	0.54
1:A:151:PHE:HB3	1:A:152:PRO:HD3	1.90	0.53
1:A:180:ASN:O	1:A:183:VAL:N	2.42	0.53
1:A:388:ARG:CD	5:A:644:HOH:O	2.56	0.53
1:B:189:ARG:O	1:B:192:GLU:HG2	2.09	0.53
1:A:203:ARG:HG3	1:A:216:TRP:CG	2.45	0.52
1:A:288:THR:HB	1:A:311:LEU:HD11	1.92	0.52
1:B:384:ASN:HB3	1:B:387:LEU:O	2.09	0.52
1:B:388:ARG:HD3	5:B:643:HOH:O	2.10	0.51
1:B:189:ARG:NE	5:B:603:HOH:O	2.25	0.51
1:A:203:ARG:HD3	5:A:791:HOH:O	2.11	0.51
1:B:348:GLN:NE2	4:B:503:GOL:H2	2.26	0.51
2:B:501:HEM:HMB1	2:B:501:HEM:HBB2	1.93	0.51
1:B:42:ARG:CG	1:B:42:ARG:NH1	2.70	0.50
1:A:251:VAL:HG23	1:A:279:ALA:HB2	1.93	0.50
1:A:256:ALA:HA	1:A:373:LEU:HD22	1.93	0.49
1:A:273:ARG:NH2	5:A:625:HOH:O	2.43	0.49
1:A:304:ILE:HD13	1:A:310:ILE:HG21	1.93	0.49
1:A:8:LEU:HD12	1:A:9:PRO:HD2	1.94	0.48
1:A:123:GLU:HG2	1:A:369:ARG:HH22	1.77	0.48
1:A:189:ARG:HG3	1:A:192:GLU:OE1	2.13	0.48
1:A:388:ARG:HD2	5:A:644:HOH:O	2.12	0.48
1:A:51:TYR:HA	1:A:317:ALA:HB1	1.95	0.48
1:A:324:TRP:HB2	1:A:327:PRO:HB3	1.96	0.47
1:B:288:THR:HB	1:B:311:LEU:HD11	1.95	0.47
1:A:37:VAL:HG22	1:A:48:LEU:HD21	1.97	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:256:ALA:HB2	1:A:373:LEU:CD2	2.45	0.47
1:B:10:VAL:HG21	1:B:41:ARG:NH1	2.30	0.47
1:B:157:ILE:O	1:B:162:ARG:NH2	2.46	0.47
1:A:285:PRO:HD2	2:A:501:HEM:HBC1	1.97	0.47
1:B:189:ARG:NH2	5:B:603:HOH:O	2.47	0.46
1:B:209:ASP:HA	5:B:601:HOH:O	2.15	0.46
1:A:388:ARG:HD3	5:A:644:HOH:O	2.15	0.46
1:B:206:LEU:HD13	1:B:212:GLY:HA3	1.98	0.45
1:B:50:ARG:HB2	1:B:53:GLN:OE1	2.17	0.45
1:B:273:ARG:HH12	1:B:356:ARG:HD3	1.81	0.45
2:B:501:HEM:HMC1	2:B:501:HEM:HBC2	1.97	0.45
1:A:186:GLY:C	1:A:188:PRO:HD2	2.37	0.45
1:B:86[A]:ILE:O	1:B:86[A]:ILE:CG2	2.64	0.45
1:A:297:THR:O	1:A:303:THR:HA	2.16	0.45
1:A:106:LEU:HD11	1:A:236:LEU:HD21	1.99	0.44
1:A:50:ARG:HB2	1:A:53:GLN:OE1	2.18	0.44
1:A:351:GLY:HA3	2:A:501:HEM:C2B	2.53	0.44
1:A:288:THR:HA	1:A:312:MET:O	2.18	0.44
1:B:349:CYS:HA	2:B:501:HEM:CHA	2.48	0.44
1:B:240:GLY:HA2	2:B:501:HEM:C3B	2.52	0.44
1:B:283:GLU:HA	5:B:608:HOH:O	2.18	0.43
4:B:504:GOL:H31	5:B:803:HOH:O	2.18	0.43
1:A:349:CYS:HA	2:A:501:HEM:CHA	2.49	0.42
1:A:280:VAL:O	1:A:284:SER:HA	2.20	0.42
1:B:280:VAL:HG12	1:B:341:GLY:HA3	2.02	0.42
1:A:39:TYR:CE1	1:A:44:ASP:HA	2.54	0.42
1:B:273:ARG:HD3	1:B:360:GLU:OE1	2.19	0.42
1:A:278:GLU:OE1	1:A:334:ARG:HD2	2.19	0.42
1:B:304:ILE:HD13	1:B:310:ILE:HG21	2.01	0.42
1:A:290:PHE:HA	1:A:310:ILE:O	2.20	0.41
1:A:187:ALA:H	1:A:190:VAL:HG23	1.85	0.41
1:B:243:THR:HB	2:B:501:HEM:C2C	2.56	0.41
1:A:256:ALA:CA	1:A:373:LEU:HD22	2.51	0.40
1:B:207:THR:HG23	5:B:648:HOH:O	2.21	0.40
1:B:80:PRO:HB2	1:B:82:ARG:O	2.22	0.40
1:B:239:ALA:HB2	3:B:502:TWO:C6	2.51	0.40

All (2) 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)
5:A:754:HOH:O	5:B:603:HOH:O[1_656]	2.09	0.11
5:A:623:HOH:O	5:B:826:HOH:O[1_556]	2.18	0.02

### 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	393/420 (94%)	378 (96%)	14 (4%)	1 (0%)	37	25
1	B	395/420 (94%)	384 (97%)	10 (2%)	1 (0%)	37	25
All	All	788/840 (94%)	762 (97%)	24 (3%)	2 (0%)	37	25

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	187	ALA
1	B	139	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	307/330 (93%)	301 (98%)	6 (2%)	50	37
1	B	309/330 (94%)	298 (96%)	11 (4%)	30	15
All	All	616/660 (93%)	599 (97%)	17 (3%)	42	24

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	SER
1	A	112[A]	GLN
1	A	112[B]	GLN
1	A	128	GLN
1	A	227	GLU
1	A	277	ASP
1	B	25	LEU
1	B	42	ARG
1	B	86[A]	ILE
1	B	86[B]	ILE
1	B	109	ARG
1	B	112	GLN
1	B	242	ASP
1	B	264	ARG
1	B	277	ASP
1	B	283	GLU
1	B	399	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	94	HIS

### 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 [i](#)

6 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	TWO	B	502	-	13,13,13	0.89	1 (7%)	17,17,17	0.91	0
2	HEM	A	501	1	41,50,50	1.31	4 (9%)	45,82,82	1.94	13 (28%)
4	GOL	B	503	-	5,5,5	0.15	0	5,5,5	0.27	0
4	GOL	B	504	-	5,5,5	0.14	0	5,5,5	0.47	0
2	HEM	B	501	1	41,50,50	1.29	4 (9%)	45,82,82	1.96	14 (31%)
3	TWO	A	502	-	13,13,13	0.74	0	17,17,17	0.77	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	TWO	B	502	-	-	0/8/8/8	0/1/1/1
2	HEM	A	501	1	-	4/12/54/54	-
4	GOL	B	503	-	-	2/4/4/4	-
4	GOL	B	504	-	-	4/4/4/4	-
2	HEM	B	501	1	-	0/12/54/54	-
3	TWO	A	502	-	-	0/8/8/8	0/1/1/1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	HEM	C1B-NB	-4.56	1.32	1.40
2	B	501	HEM	C1B-NB	-3.39	1.34	1.40
2	A	501	HEM	C4B-NB	-3.36	1.31	1.38
2	B	501	HEM	C4B-NB	-2.76	1.33	1.38
2	A	501	HEM	C4D-ND	-2.56	1.35	1.40
2	B	501	HEM	C4D-ND	-2.33	1.36	1.40
3	B	502	TWO	O2-C7	-2.10	1.24	1.30
2	B	501	HEM	CHB-C1B	2.09	1.40	1.35
2	A	501	HEM	FE-NB	2.05	2.07	1.96

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	HEM	CHC-C4B-NB	5.39	130.29	124.43
2	A	501	HEM	C1B-NB-C4B	5.30	110.54	105.07
2	B	501	HEM	CHC-C4B-NB	4.14	128.93	124.43
2	B	501	HEM	C1B-NB-C4B	3.96	109.16	105.07
2	B	501	HEM	CHD-C1D-ND	3.72	128.48	124.43
2	B	501	HEM	CAD-C3D-C4D	3.44	130.68	124.66
2	B	501	HEM	CBD-CAD-C3D	3.41	122.11	112.63
2	B	501	HEM	CHD-C1D-C2D	-3.39	119.69	124.98
2	A	501	HEM	CHA-C4D-C3D	-3.19	119.35	125.33
2	A	501	HEM	CBD-CAD-C3D	2.73	120.21	112.63
2	A	501	HEM	CHA-C4D-ND	2.67	127.69	124.38
2	B	501	HEM	CMC-C2C-C3C	2.62	129.58	124.68
2	A	501	HEM	O2A-CGA-CBA	2.56	122.26	114.03
2	B	501	HEM	CHA-C4D-C3D	-2.55	120.54	125.33
2	A	501	HEM	O1A-CGA-CBA	-2.51	115.03	123.08
2	B	501	HEM	C4B-CHC-C1C	2.50	125.86	122.56
2	B	501	HEM	O2A-CGA-CBA	2.42	121.81	114.03
2	B	501	HEM	O2D-CGD-O1D	-2.41	117.30	123.30
2	A	501	HEM	CMA-C3A-C4A	-2.36	124.83	128.46
2	A	501	HEM	CHD-C1D-C2D	-2.28	121.42	124.98
2	A	501	HEM	C3B-C2B-C1B	-2.27	104.80	106.49
2	A	501	HEM	C3D-C4D-ND	2.25	112.67	110.17
2	A	501	HEM	O2D-CGD-CBD	2.20	121.09	114.03
2	A	501	HEM	CMB-C2B-C1B	2.13	128.28	125.04
2	B	501	HEM	CMD-C2D-C1D	2.12	128.27	125.04
2	B	501	HEM	C4B-C3B-C2B	-2.06	105.48	107.11
2	B	501	HEM	C3D-C4D-ND	2.02	112.42	110.17

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	504	GOL	C1-C2-C3-O3
4	B	504	GOL	O2-C2-C3-O3
4	B	503	GOL	O1-C1-C2-C3
4	B	504	GOL	O1-C1-C2-C3
4	B	503	GOL	O1-C1-C2-O2
4	B	504	GOL	O1-C1-C2-O2
2	A	501	HEM	CAD-CBD-CGD-O2D
2	A	501	HEM	CAA-CBA-CGA-O2A
2	A	501	HEM	CAA-CBA-CGA-O1A

*Continued on next page...*



*Continued from previous page...*

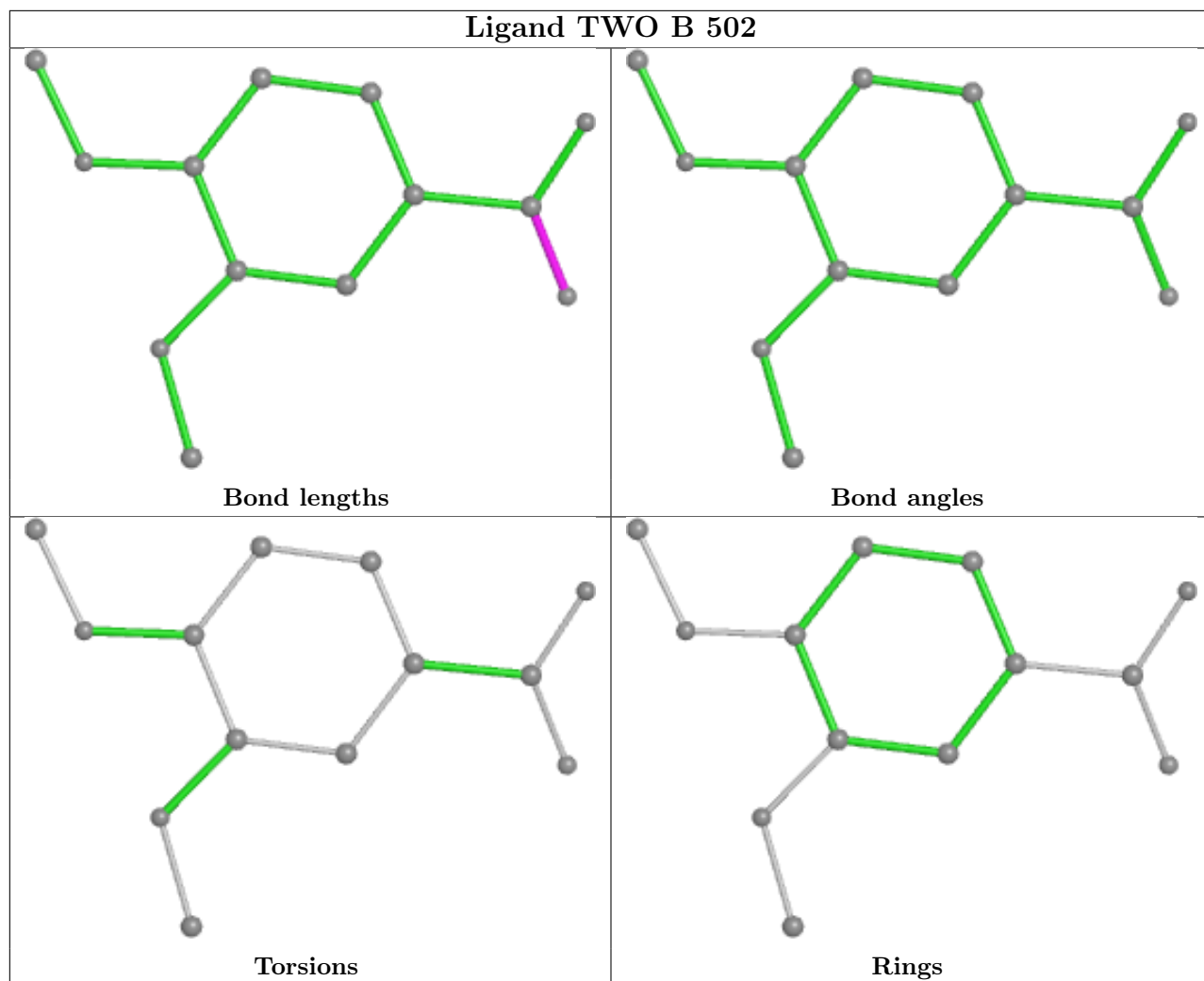
Mol	Chain	Res	Type	Atoms
2	A	501	HEM	CAD-CBD-CGD-O1D

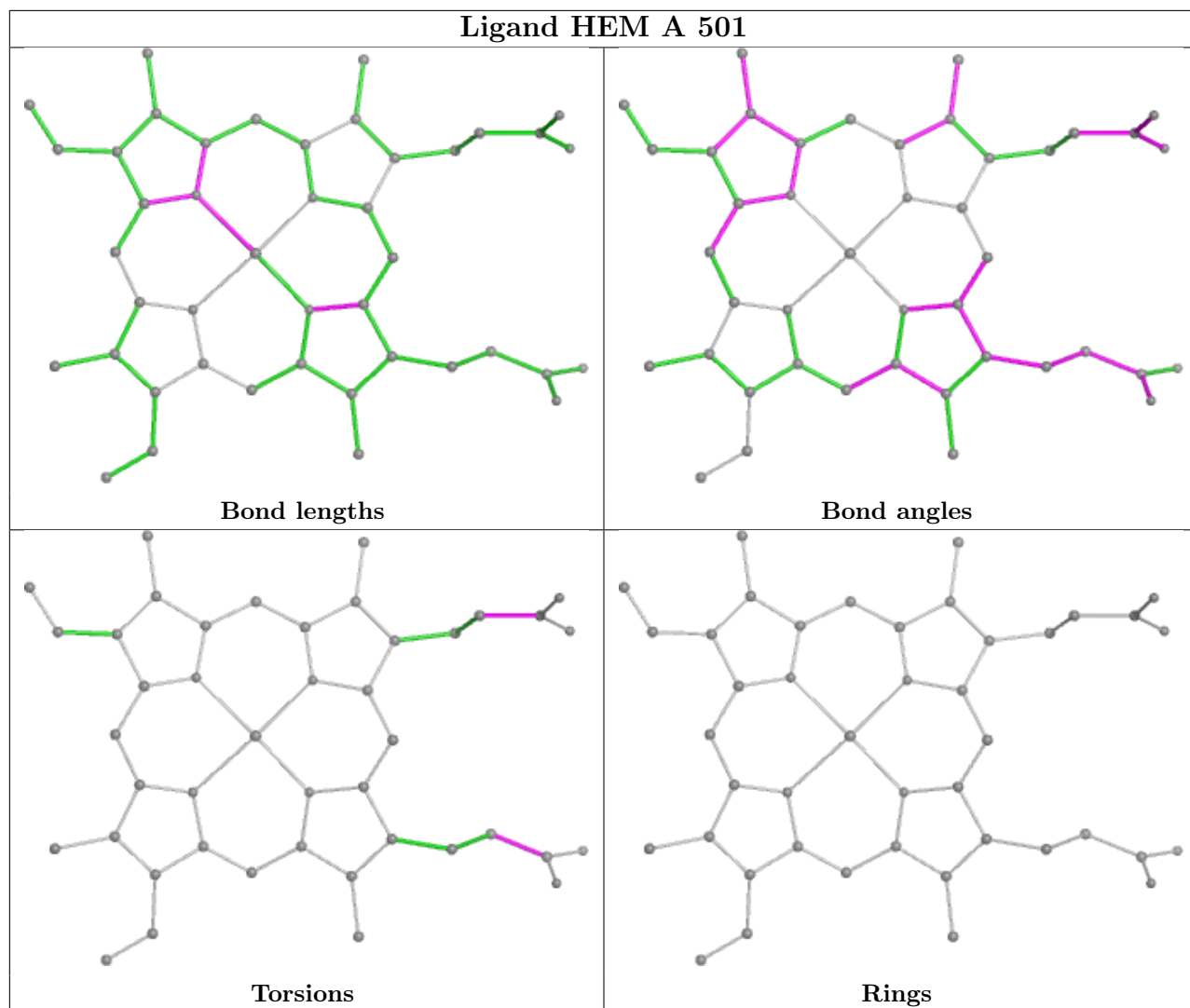
There are no ring outliers.

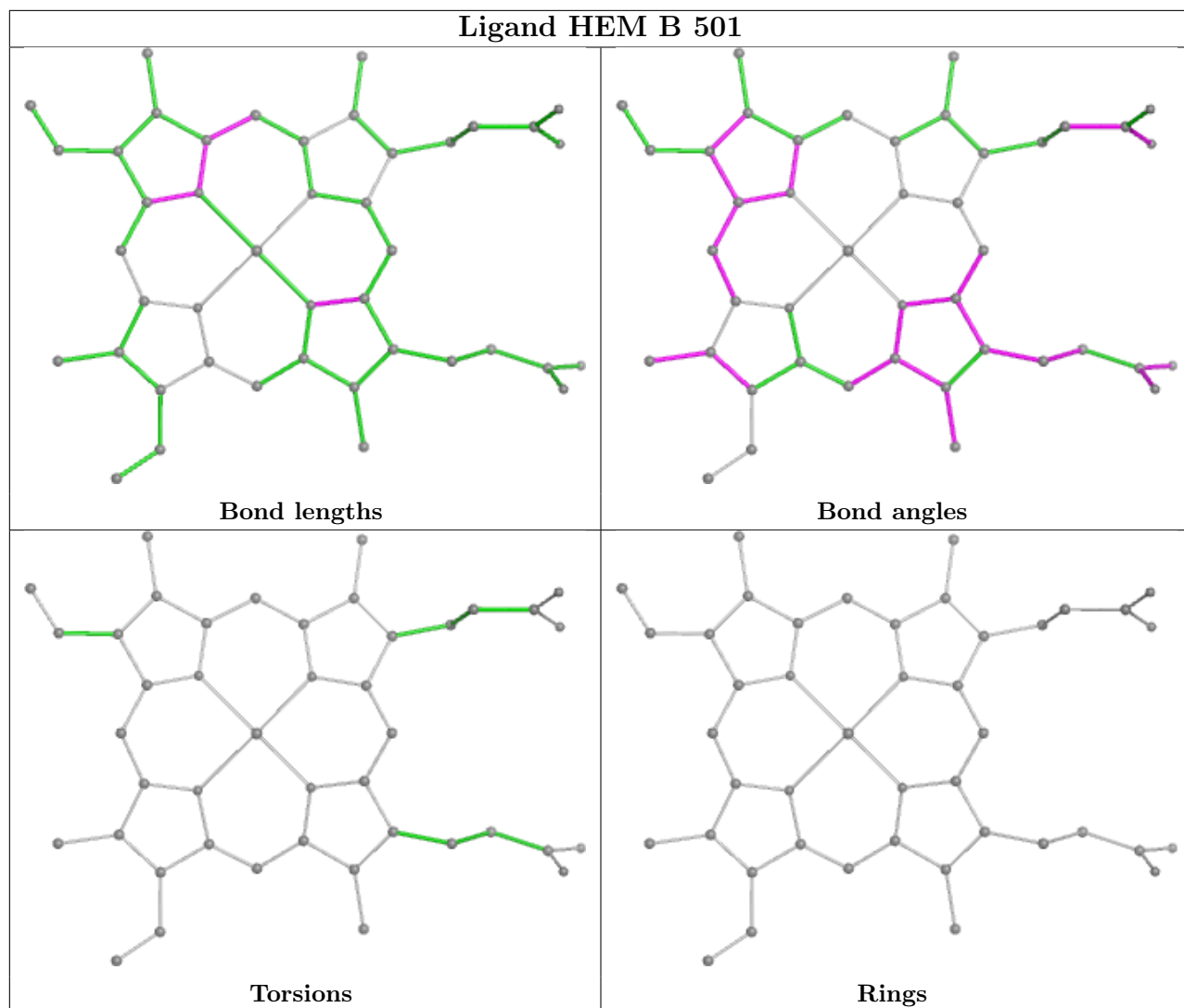
5 monomers are involved in 15 short contacts:

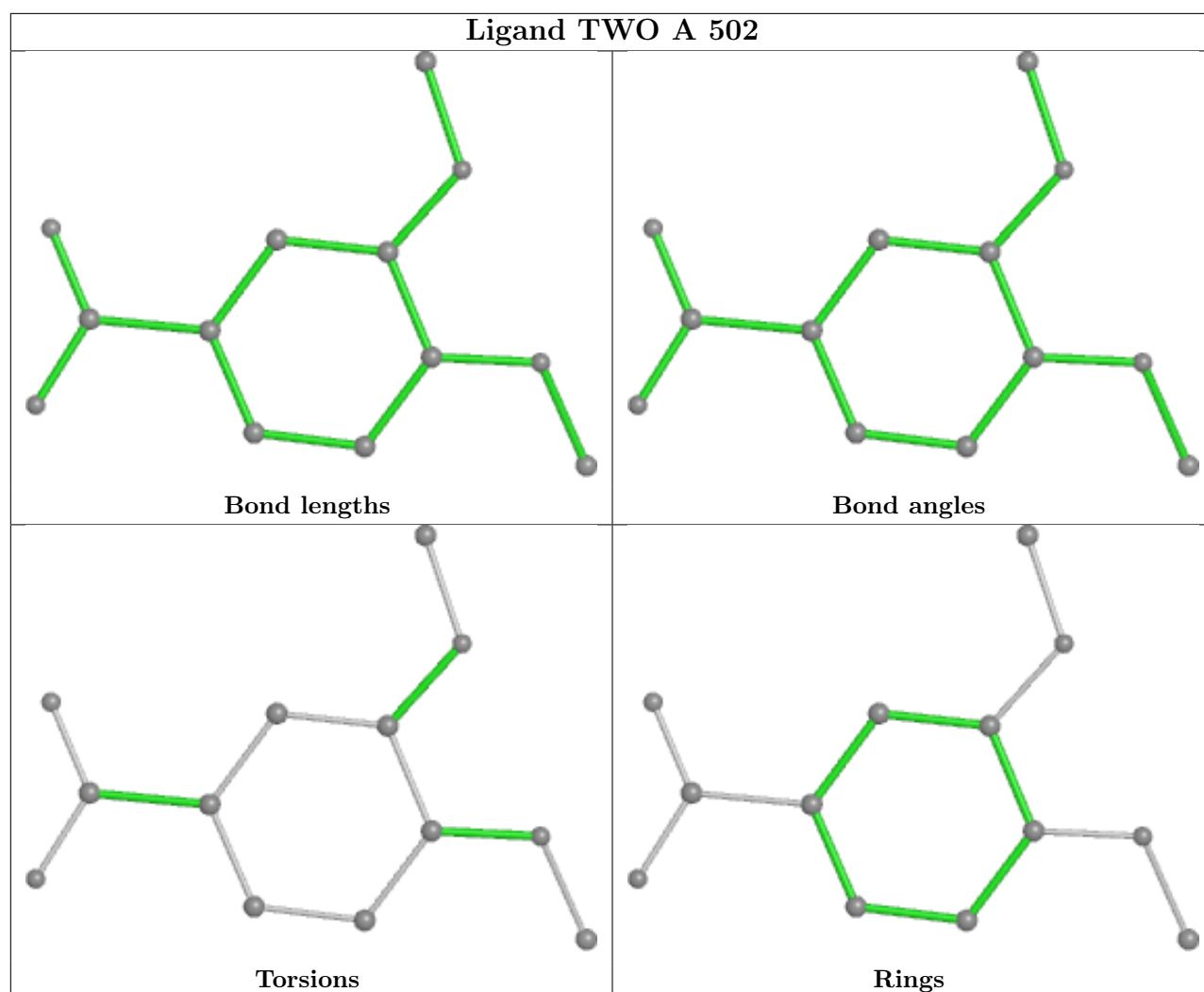
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	502	TWO	1	0
2	A	501	HEM	6	0
4	B	503	GOL	1	0
4	B	504	GOL	1	0
2	B	501	HEM	6	0

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

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	393/420 (93%)	0.55	29 (7%) 22 23	15, 23, 45, 73	1 (0%)
1	B	393/420 (93%)	0.48	13 (3%) 49 53	9, 22, 38, 51	2 (0%)
All	All	786/840 (93%)	0.52	42 (5%) 33 35	9, 23, 40, 73	3 (0%)

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	183	VAL	6.1
1	A	185	LYS	5.0
1	A	187	ALA	4.9
1	A	184	GLU	4.5
1	A	182	LEU	4.3
1	A	302	ALA	4.0
1	A	181	GLY	3.9
1	B	160	ALA	3.8
1	A	299	ILE	3.6
1	A	188	PRO	3.6
1	B	390	TRP	3.3
1	B	207	THR	3.0
1	A	190	VAL	2.9
1	A	300	GLY	2.8
1	A	149	ARG	2.7
1	B	304	ILE	2.7
1	A	8	LEU	2.7
1	B	161	GLY	2.7
1	B	208	GLY	2.7
1	B	380[A]	HIS	2.7
1	A	240	GLY	2.6
1	A	272	ALA	2.6
1	A	373	LEU	2.6
1	A	273	ARG	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	15	PRO	2.4
1	B	209	ASP	2.4
1	A	124	ALA	2.4
1	A	207	THR	2.4
1	A	298	GLU	2.3
1	B	8	LEU	2.2
1	B	159	ASP	2.1
1	A	148	LEU	2.1
1	A	166	LEU	2.1
1	A	383	LEU	2.1
1	B	366	LEU	2.1
1	A	128	GLN	2.1
1	A	186	GLY	2.1
1	B	158	PRO	2.0
1	B	368	SER	2.0
1	A	274	THR	2.0
1	A	179	ALA	2.0
1	A	125	LEU	2.0

## 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 monosaccharides 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.

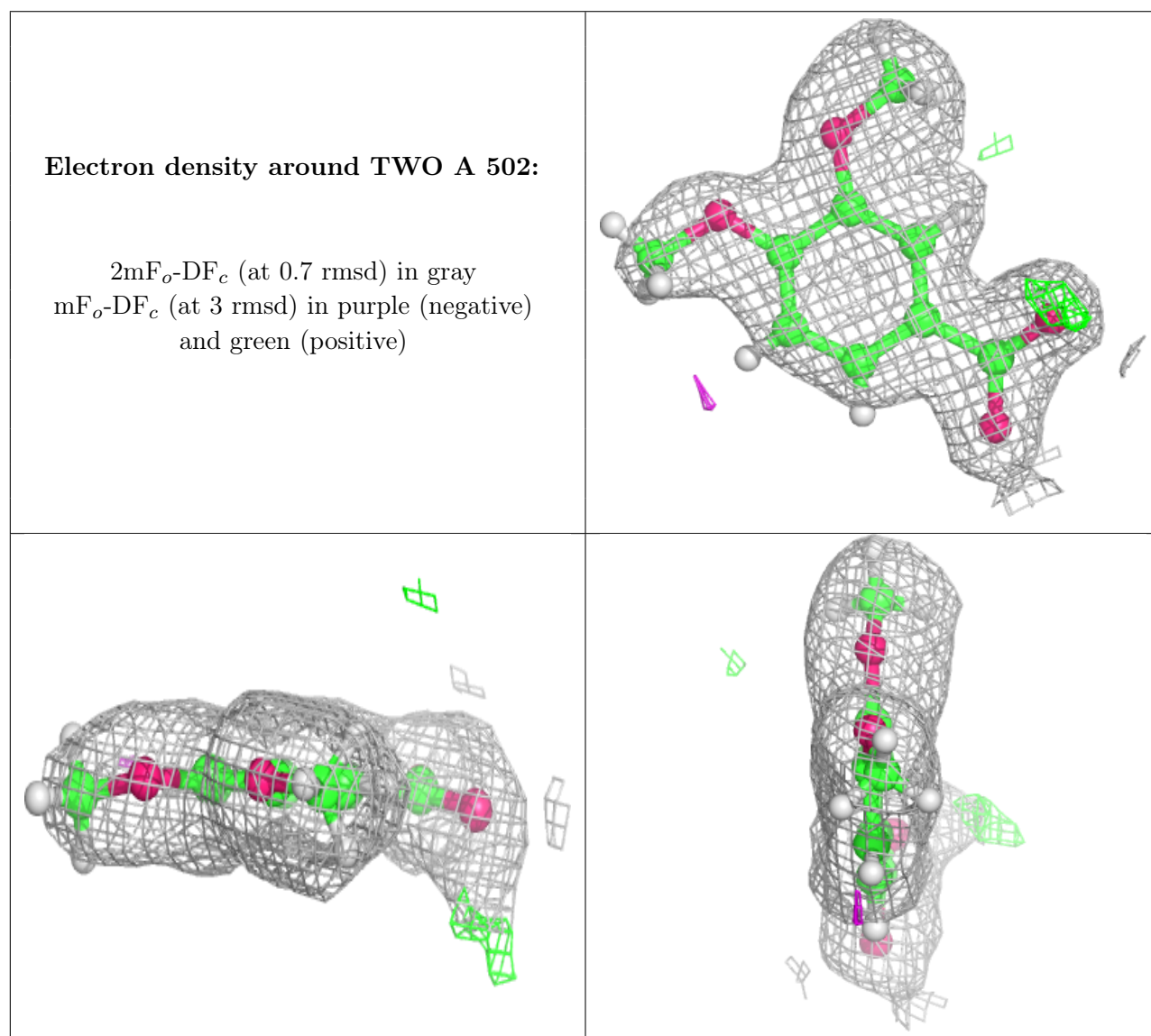
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	GOL	B	504	6/6	0.60	0.23	29,35,40,40	14
4	GOL	B	503	6/6	0.82	0.15	22,28,31,31	14
3	TWO	A	502	13/13	0.92	0.08	16,21,22,24	0
3	TWO	B	502	13/13	0.95	0.07	13,16,20,20	0
2	HEM	A	501	43/43	0.96	0.08	16,18,22,25	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	HEM	B	501	43/43	0.96	0.09	15,18,26,31	0

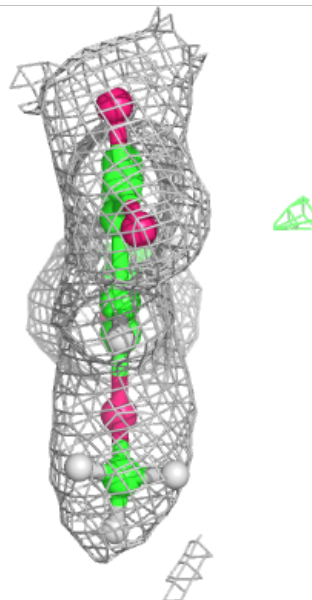
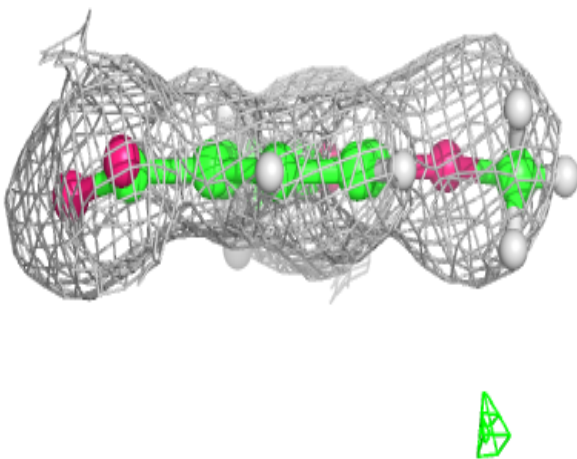
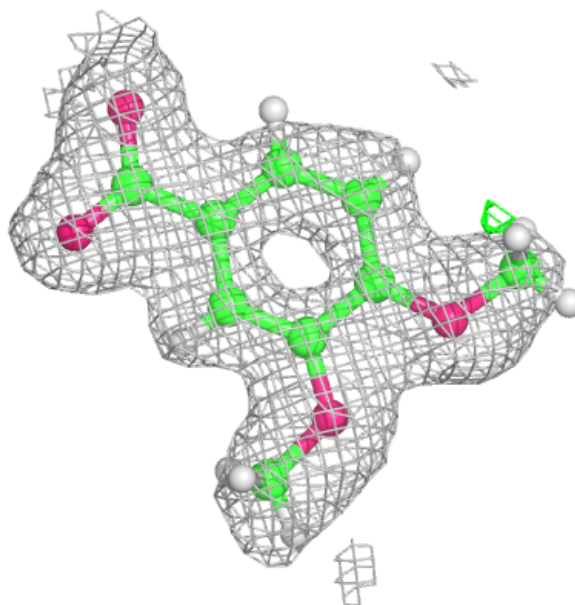
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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





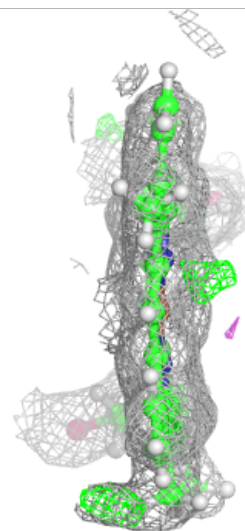
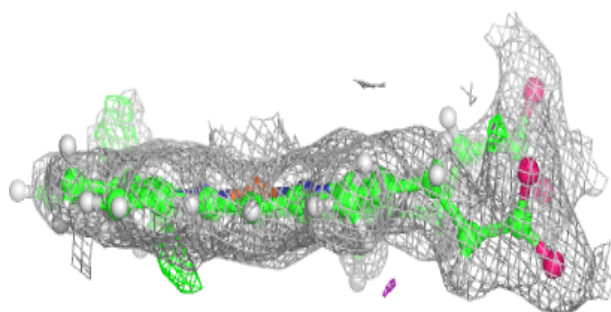
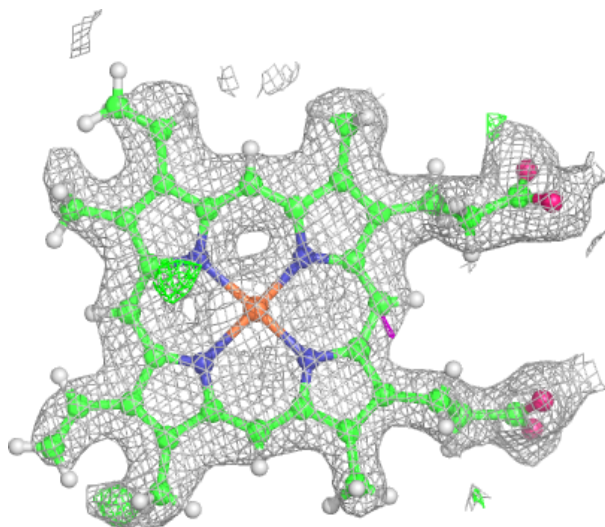
**Electron density around TWO B 502:**

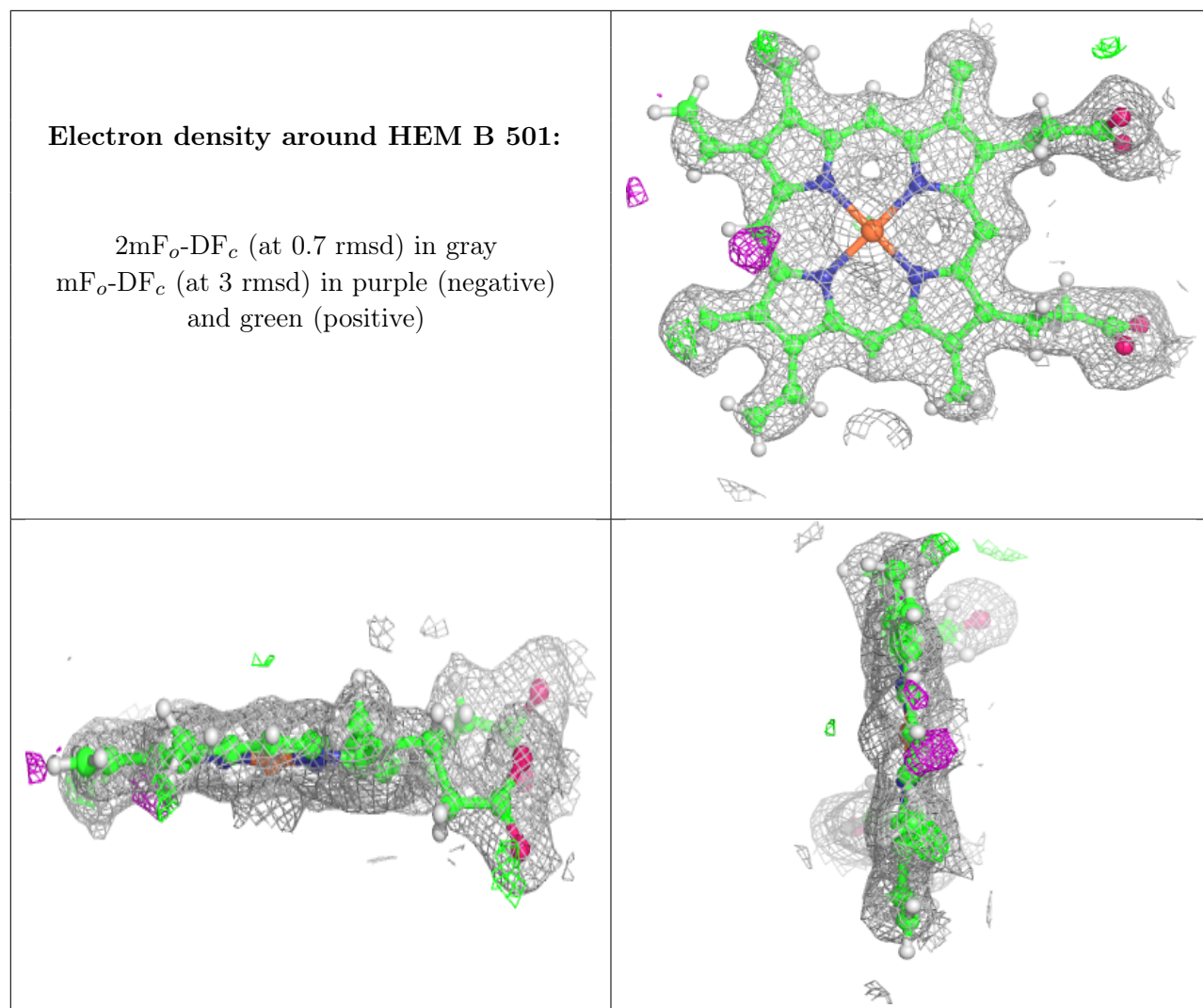
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around HEM A 501:**

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





## 6.5 Other polymers [i](#)

There are no such residues in this entry.