



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 3, 2024 – 12:12 PM EDT

PDB ID : 8U3P  
Title : 1.79 Angstrom resolution crystal structure of KatG from Mycobacterium tuberculosis with an MYW cofactor after heat incubation for 60 minutes  
Authors : Liu, A.; Li, J.; Ran, D.  
Deposited on : 2023-09-08  
Resolution : 1.79 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 : 2022.3.0, CSD as543be (2022)  
Xtrriage (Phenix) : 1.20.1  
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) : 2.38.3

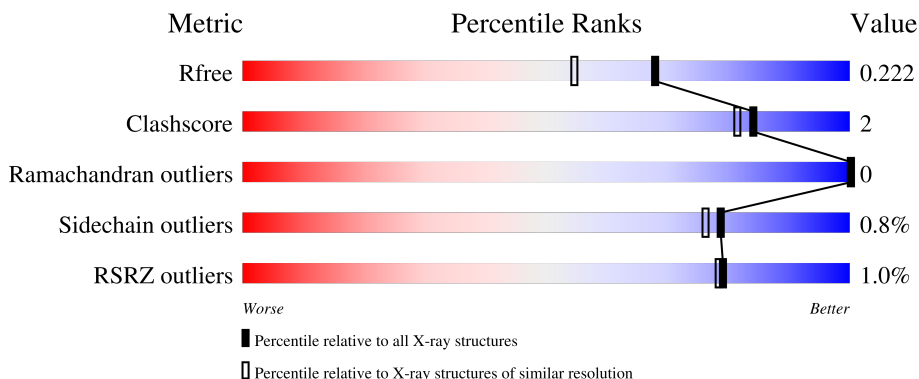
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.79 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	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

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	741	 2% 91% 5% .
1	B	741	 1% 91% 6% .
1	C	741	 1% 92% 5% .
1	D	741	 1% 92% . .

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 25857 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 Catalase-peroxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	717	5529	3515	953	1042	19	0	0	0
1	B	717	5529	3515	953	1042	19	0	0	0
1	C	717	5535	3519	954	1043	19	0	1	0
1	D	717	5529	3515	953	1042	19	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP A0A0D5ZBI4
A	1	HIS	-	expression tag	UNP A0A0D5ZBI4
B	0	GLY	-	expression tag	UNP A0A0D5ZBI4
B	1	HIS	-	expression tag	UNP A0A0D5ZBI4
C	0	GLY	-	expression tag	UNP A0A0D5ZBI4
C	1	HIS	-	expression tag	UNP A0A0D5ZBI4
D	0	GLY	-	expression tag	UNP A0A0D5ZBI4
D	1	HIS	-	expression tag	UNP A0A0D5ZBI4

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Na 1	0	0
2	B	1	Total 1	Na 1	0	0
2	C	1	Total 1	Na 1	0	0
2	D	1	Total 1	Na 1	0	0

- Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	C	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0
5	D	1	Total C O 6 3 3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	880	Total O 880 880	0	0
6	B	843	Total O 843 843	0	0
6	C	860	Total O 860 860	0	0

*Continued on next page...*

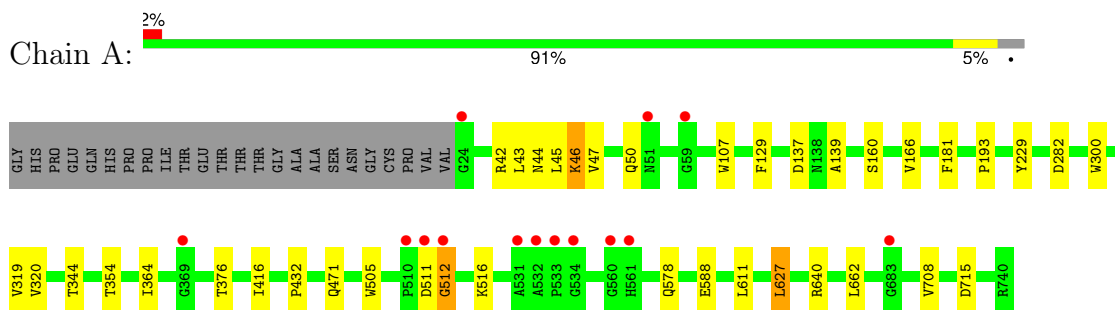
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	D	858	Total 858	O 858	0	0

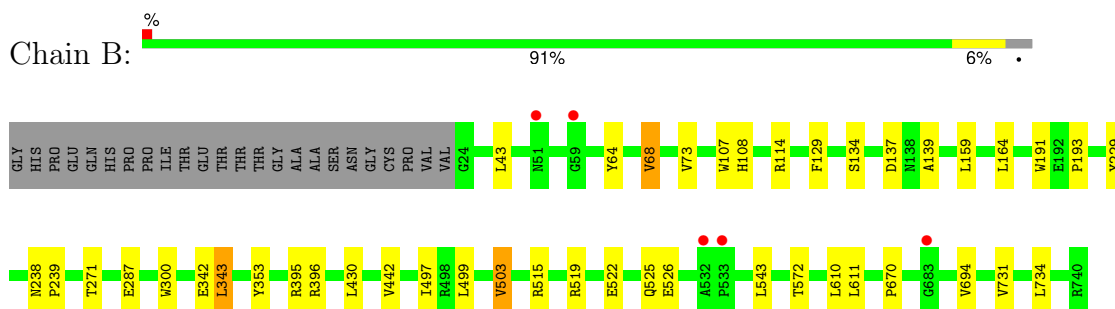
### 3 Residue-property plots [i](#)

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.

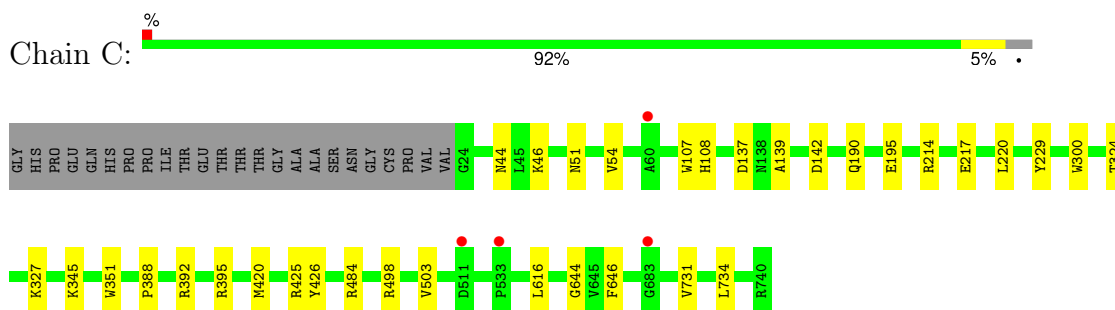
- Molecule 1: Catalase-peroxidase



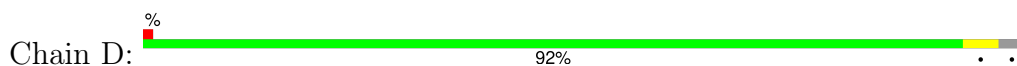
- Molecule 1: Catalase-peroxidase



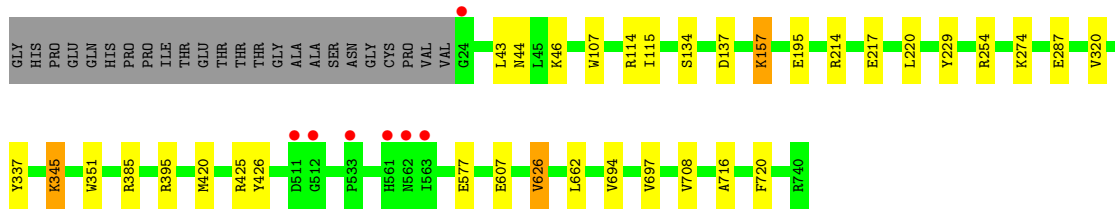
- Molecule 1: Catalase-peroxidase



- Molecule 1: Catalase-peroxidase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	150.58Å 150.58Å 311.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.55 – 1.79 49.55 – 1.79	Depositor EDS
% Data completeness (in resolution range)	95.5 (49.55-1.79) 91.3 (49.55-1.79)	Depositor EDS
$R_{merge}$	0.22	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.66 (at 1.79Å)	Xtrriage
Refinement program	PHENIX (1.12_2829-000)	Depositor
R, $R_{free}$	0.179 , 0.220 0.180 , 0.222	Depositor DCC
$R_{free}$ test set	316366 reflections (0.63%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.3	Xtrriage
Anisotropy	0.521	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 43.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	25857	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 57.35 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.3968e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: GOL, HEM, ACT, NA

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	0.40	0/5680	0.57	3/7732 (0.0%)
1	B	0.48	4/5680 (0.1%)	0.57	0/7732
1	C	0.39	0/5689	0.55	0/7744
1	D	0.39	0/5680	0.56	1/7732 (0.0%)
All	All	0.42	4/22729 (0.0%)	0.56	4/30940 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	503	VAL	CB-CG2	-5.70	1.40	1.52
1	B	342	GLU	CD-OE2	-5.66	1.19	1.25
1	B	68	VAL	CB-CG1	-5.18	1.42	1.52
1	B	271	THR	C-O	-5.03	1.13	1.23

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	640	ARG	NE-CZ-NH2	-8.56	116.02	120.30
1	D	626	VAL	CG1-CB-CG2	-7.45	98.99	110.90
1	A	512	GLY	N-CA-C	-6.19	97.62	113.10
1	A	627	LEU	CB-CG-CD1	5.71	120.70	111.00

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	5529	0	5355	29	0
1	B	5529	0	5355	26	0
1	C	5535	0	5363	18	0
1	D	5529	0	5355	25	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	43	0	30	0	0
3	B	43	0	30	0	0
3	C	43	0	30	0	0
3	D	43	0	30	0	0
4	A	4	0	3	0	0
4	B	8	0	6	0	0
4	D	4	0	3	1	0
5	A	24	0	32	1	0
5	B	30	0	38	5	0
5	C	12	0	16	0	0
5	D	36	0	48	4	0
6	A	880	0	0	6	0
6	B	843	0	0	3	0
6	C	860	0	0	1	0
6	D	858	0	0	0	0
All	All	25857	0	21694	99	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 2.

All (99) 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:511:ASP:HB2	1:A:512:GLY:HA2	1.42	1.02
1:D:44:ASN:ND2	1:D:46:LYS:HE2	2.01	0.74
1:A:44:ASN:OD1	1:A:46:LYS:HG2	1.89	0.71
1:D:115:ILE:HD11	5:D:1009:GOL:H31	1.71	0.70
1:A:46:LYS:HD2	6:A:1738:HOH:O	1.95	0.66
1:A:516:LYS:NZ	6:A:1103:HOH:O	2.28	0.66
1:A:511:ASP:CB	1:A:512:GLY:HA2	2.23	0.66
1:A:44:ASN:ND2	1:A:46:LYS:HG2	2.11	0.65

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:44:ASN:HD21	1:D:46:LYS:HE2	1.60	0.65
1:A:44:ASN:CG	1:A:46:LYS:HG2	2.21	0.61
1:D:46:LYS:HE3	1:D:195:GLU:OE2	2.01	0.60
1:A:511:ASP:HB2	1:A:512:GLY:CA	2.24	0.59
1:D:44:ASN:HD21	1:D:46:LYS:HG2	1.69	0.58
1:A:43:LEU:HG	1:A:611:LEU:HD22	1.85	0.58
1:A:578:GLN:NE2	6:A:1102:HOH:O	2.37	0.57
1:C:345:LYS:HE2	1:C:351:TRP:CZ3	2.39	0.56
1:B:129:PHE:HB3	1:B:193:PRO:HG3	1.88	0.56
1:A:44:ASN:HD21	1:A:46:LYS:HG2	1.69	0.55
1:D:626:VAL:HG13	1:D:720:PHE:CE1	2.41	0.55
1:A:129:PHE:HB3	1:A:193:PRO:HG3	1.88	0.55
1:B:107:TRP:HZ3	1:B:229:TYR:HH	1.55	0.55
1:A:107:TRP:HZ3	1:A:229:TYR:HH	1.52	0.55
1:D:345:LYS:HE3	1:D:351:TRP:CE3	2.43	0.54
1:D:254:ARG:HG2	5:D:1009:GOL:H11	1.90	0.54
1:D:577:GLU:H	1:D:577:GLU:CD	2.11	0.53
1:D:626:VAL:HG13	1:D:720:PHE:CD1	2.43	0.53
1:C:44:ASN:HD21	1:C:46:LYS:HB3	1.74	0.52
1:D:44:ASN:HD21	1:D:46:LYS:CE	2.23	0.52
1:B:107:TRP:CD1	1:B:108:HIS:HD2	2.27	0.52
1:B:499:LEU:O	1:B:503:VAL:HG23	2.09	0.52
1:B:43:LEU:HG	1:B:611:LEU:HD22	1.92	0.51
5:B:1008:GOL:O3	6:B:1101:HOH:O	2.19	0.51
1:D:217:GLU:HG2	1:D:220:LEU:HD12	1.93	0.51
1:D:44:ASN:HD21	1:D:46:LYS:CG	2.23	0.50
1:A:160:SER:HA	5:A:1004:GOL:H12	1.93	0.49
1:B:522:GLU:O	1:B:526:GLU:HG2	2.12	0.49
5:B:1007:GOL:H2	5:B:1009:GOL:H32	1.95	0.48
1:B:139:ALA:HA	1:B:300:TRP:CZ3	2.49	0.48
1:C:498:ARG:C	1:C:503:VAL:HG13	2.34	0.48
1:B:343:LEU:HD13	1:B:353:TYR:CZ	2.50	0.47
1:B:191:TRP:CH2	5:B:1006:GOL:H32	2.50	0.47
1:B:43:LEU:HD22	5:B:1007:GOL:H11	1.96	0.47
1:A:662:LEU:HD11	1:A:708:VAL:HG12	1.96	0.47
1:B:610:LEU:HD22	1:B:694:VAL:HG13	1.97	0.46
1:C:324:THR:HB	1:C:327:LYS:HB3	1.97	0.46
1:D:43:LEU:HD22	5:D:1007:GOL:H32	1.98	0.46
1:B:64:TYR:CZ	1:B:68:VAL:CG1	2.99	0.46
1:A:139:ALA:HA	1:A:300:TRP:CZ3	2.51	0.45
1:A:319:VAL:HG22	1:A:376:THR:HB	1.98	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:715:ASP:OD2	1:D:157:LYS:HG2	2.16	0.45
1:B:64:TYR:CE1	1:B:68:VAL:HG13	2.51	0.45
1:C:51:ASN:ND2	1:C:190:GLN:HG2	2.31	0.45
1:A:46:LYS:HG3	1:A:47:VAL:N	2.30	0.45
1:D:607:GLU:HG3	5:D:1007:GOL:H11	1.98	0.45
1:D:716:ALA:HB2	4:D:1003:ACT:H3	1.98	0.45
1:B:670:PRO:HB3	5:B:1008:GOL:H31	1.98	0.45
1:B:395:ARG:NH1	6:B:1106:HOH:O	2.38	0.44
1:C:107:TRP:HZ3	1:C:229:TYR:HH	1.60	0.44
1:D:337:TYR:CE1	1:D:385:ARG:HG3	2.51	0.44
1:C:731:VAL:HA	1:C:734:LEU:HG	1.99	0.44
1:A:344:THR:HG21	1:A:354:THR:OG1	2.17	0.44
1:B:497:ILE:O	1:B:503:VAL:HG22	2.18	0.44
1:A:50:GLN:NE2	6:A:1145:HOH:O	2.51	0.44
1:A:181:PHE:CE2	1:A:432:PRO:HG2	2.53	0.44
1:A:505:TRP:CE2	1:A:588:GLU:HB2	2.53	0.44
1:C:54:VAL:HG23	6:C:3091:HOH:O	2.18	0.43
1:D:107:TRP:HZ3	1:D:229:TYR:HH	1.62	0.43
1:A:471:GLN:HE22	1:A:516:LYS:NZ	2.16	0.43
1:D:134:SER:HB3	1:D:287:GLU:HG3	2.01	0.43
1:D:420:MET:O	1:D:425:ARG:HD3	2.18	0.43
1:C:46:LYS:HE3	1:C:195:GLU:HG3	2.01	0.43
1:D:662:LEU:HD11	1:D:708:VAL:HG12	2.01	0.43
1:C:498:ARG:HA	1:C:503:VAL:CG1	2.49	0.43
1:D:694:VAL:O	1:D:697:VAL:HG12	2.19	0.43
1:A:320:VAL:HB	1:A:364:ILE:HG21	2.01	0.42
1:C:217:GLU:HG2	1:C:220:LEU:HD12	2.02	0.42
1:D:44:ASN:HD21	1:D:46:LYS:CD	2.32	0.42
1:A:166:VAL:HG21	1:A:416:ILE:HB	2.01	0.42
1:C:420:MET:O	1:C:425:ARG:HD3	2.19	0.42
1:B:64:TYR:CZ	1:B:68:VAL:HG13	2.54	0.42
1:C:139:ALA:HA	1:C:300:TRP:CZ3	2.54	0.42
1:B:731:VAL:HA	1:B:734:LEU:HG	2.02	0.42
1:B:442:VAL:HG13	1:B:572:THR:OG1	2.20	0.41
1:C:142:ASP:N	1:C:142:ASP:OD1	2.53	0.41
1:B:515:ARG:O	1:B:519:ARG:HG3	2.21	0.41
1:B:73:VAL:HG11	1:B:430:LEU:HD11	2.02	0.41
1:B:519:ARG:HG2	6:B:1266:HOH:O	2.21	0.41
1:A:42:ARG:NH1	6:A:1163:HOH:O	2.54	0.41
1:A:282:ASP:HB3	6:A:1771:HOH:O	2.21	0.41
1:C:484:ARG:HB2	1:C:616:LEU:HD22	2.03	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:388:PRO:O	1:C:392:ARG:HG3	2.20	0.41
1:A:45:LEU:CD2	1:A:611:LEU:HD21	2.51	0.41
1:B:134:SER:HB3	1:B:287:GLU:HG3	2.03	0.41
1:B:525:GLN:HB2	1:B:543:LEU:HD12	2.01	0.41
1:C:107:TRP:CD1	1:C:108:HIS:HD2	2.39	0.41
1:C:644:GLY:HA2	1:C:646:PHE:CZ	2.56	0.41
1:B:238:ASN:HA	1:B:239:PRO:HD3	1.95	0.40
1:B:159:LEU:HD21	1:B:164:LEU:HD13	2.04	0.40
1:D:274:LYS:HB2	1:D:320:VAL:HG22	2.03	0.40

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	715/741 (96%)	700 (98%)	15 (2%)	0	100	100
1	B	715/741 (96%)	701 (98%)	14 (2%)	0	100	100
1	C	716/741 (97%)	701 (98%)	15 (2%)	0	100	100
1	D	715/741 (96%)	706 (99%)	9 (1%)	0	100	100
All	All	2861/2964 (96%)	2808 (98%)	53 (2%)	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	565/584 (97%)	562 (100%)	3 (0%)	86	86
1	B	565/584 (97%)	561 (99%)	4 (1%)	81	79
1	C	566/584 (97%)	562 (99%)	4 (1%)	81	79
1	D	565/584 (97%)	558 (99%)	7 (1%)	67	62
All	All	2261/2336 (97%)	2243 (99%)	18 (1%)	79	76

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	46	LYS
1	A	137	ASP
1	A	627	LEU
1	B	114	ARG
1	B	137	ASP
1	B	343	LEU
1	B	396	ARG
1	C	137	ASP
1	C	214	ARG
1	C	395	ARG
1	C	426	TYR
1	D	114	ARG
1	D	137	ASP
1	D	157	LYS
1	D	214	ARG
1	D	345	LYS
1	D	395	ARG
1	D	426	TYR

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

Mol	Chain	Res	Type
1	A	471	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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 29 ligands modelled in this entry, 4 are monoatomic - leaving 25 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
5	GOL	C	1004	-	5,5,5	0.81	0	5,5,5	1.13	1 (20%)
5	GOL	D	1008	-	5,5,5	0.85	0	5,5,5	1.20	0
5	GOL	B	1007	-	5,5,5	0.98	0	5,5,5	1.41	1 (20%)
5	GOL	A	1004	-	5,5,5	1.01	0	5,5,5	0.86	0
5	GOL	A	1007	-	5,5,5	0.88	0	5,5,5	1.14	0
5	GOL	D	1007	-	5,5,5	0.86	0	5,5,5	1.21	1 (20%)
4	ACT	B	1003	-	3,3,3	1.24	0	3,3,3	1.28	0
3	HEM	B	1002	1	42,50,50	1.51	5 (11%)	46,82,82	1.47	7 (15%)
4	ACT	B	1004	-	3,3,3	1.30	0	3,3,3	1.35	0
5	GOL	B	1006	-	5,5,5	0.99	0	5,5,5	1.40	1 (20%)
4	ACT	A	1003	-	3,3,3	1.48	1 (33%)	3,3,3	1.51	0
3	HEM	A	1002	1	42,50,50	1.55	7 (16%)	46,82,82	1.39	6 (13%)
4	ACT	D	1003	-	3,3,3	1.28	0	3,3,3	1.46	0
5	GOL	A	1005	-	5,5,5	0.84	0	5,5,5	1.12	0
5	GOL	D	1009	-	5,5,5	0.50	0	5,5,5	1.22	0
5	GOL	B	1008	-	5,5,5	0.87	0	5,5,5	1.20	1 (20%)
5	GOL	D	1004	-	5,5,5	0.87	0	5,5,5	1.07	0
3	HEM	D	1002	1,6	42,50,50	1.62	7 (16%)	46,82,82	1.37	3 (6%)
5	GOL	B	1009	-	5,5,5	1.07	0	5,5,5	1.27	1 (20%)
5	GOL	C	1003	-	5,5,5	1.03	0	5,5,5	0.77	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	GOL	D	1005	-	5,5,5	1.42	1 (20%)	5,5,5	0.86	0
5	GOL	B	1005	-	5,5,5	0.91	0	5,5,5	1.29	1 (20%)
3	HEM	C	1002	1,6	42,50,50	1.56	6 (14%)	46,82,82	1.48	6 (13%)
5	GOL	A	1006	-	5,5,5	0.93	0	5,5,5	1.18	1 (20%)
5	GOL	D	1006	-	5,5,5	0.97	0	5,5,5	1.06	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
5	GOL	C	1004	-	-	0/4/4/4	-
5	GOL	D	1008	-	-	1/4/4/4	-
5	GOL	B	1007	-	-	0/4/4/4	-
5	GOL	A	1004	-	-	1/4/4/4	-
5	GOL	A	1007	-	-	2/4/4/4	-
5	GOL	D	1007	-	-	2/4/4/4	-
3	HEM	B	1002	1	-	2/12/54/54	-
5	GOL	B	1006	-	-	2/4/4/4	-
3	HEM	A	1002	1	-	2/12/54/54	-
5	GOL	A	1005	-	-	0/4/4/4	-
5	GOL	D	1009	-	-	2/4/4/4	-
5	GOL	B	1008	-	-	2/4/4/4	-
5	GOL	D	1004	-	-	1/4/4/4	-
3	HEM	D	1002	1,6	-	2/12/54/54	-
5	GOL	B	1009	-	-	2/4/4/4	-
5	GOL	C	1003	-	-	0/4/4/4	-
5	GOL	D	1005	-	-	0/4/4/4	-
5	GOL	B	1005	-	-	0/4/4/4	-
3	HEM	C	1002	1,6	-	3/12/54/54	-
5	GOL	A	1006	-	-	2/4/4/4	-
5	GOL	D	1006	-	-	1/4/4/4	-

All (27) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1002	HEM	C3C-C2C	-4.40	1.34	1.40
3	D	1002	HEM	C3C-C2C	-4.17	1.34	1.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	1002	HEM	C3C-CAC	3.78	1.56	1.47
3	B	1002	HEM	C3C-C2C	-3.67	1.35	1.40
3	C	1002	HEM	C3C-C2C	-3.55	1.35	1.40
3	D	1002	HEM	C3C-CAC	3.54	1.55	1.47
3	B	1002	HEM	C3C-CAC	3.53	1.55	1.47
3	A	1002	HEM	C3C-CAC	3.47	1.55	1.47
3	C	1002	HEM	C3C-C4C	3.35	1.46	1.41
3	C	1002	HEM	CAB-C3B	3.12	1.55	1.47
3	A	1002	HEM	CAB-C3B	3.11	1.55	1.47
3	B	1002	HEM	C3C-C4C	3.07	1.45	1.41
3	D	1002	HEM	CAB-C3B	3.04	1.55	1.47
3	B	1002	HEM	CAB-C3B	2.97	1.55	1.47
3	A	1002	HEM	C3C-C4C	2.89	1.45	1.41
3	D	1002	HEM	FE-ND	2.69	2.13	1.98
3	C	1002	HEM	CMD-C2D	2.44	1.55	1.50
3	D	1002	HEM	CMB-C2B	2.43	1.55	1.50
3	D	1002	HEM	CMD-C2D	2.34	1.55	1.50
3	A	1002	HEM	FE-ND	2.33	2.11	1.98
3	A	1002	HEM	CMD-C2D	2.24	1.55	1.50
3	D	1002	HEM	C3C-C4C	2.23	1.44	1.41
3	A	1002	HEM	CMB-C2B	2.22	1.55	1.50
3	B	1002	HEM	CMB-C2B	2.11	1.55	1.50
5	D	1005	GOL	C3-C2	2.08	1.59	1.51
3	C	1002	HEM	FE-ND	2.06	2.09	1.98
4	A	1003	ACT	CH3-C	2.03	1.57	1.49

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1002	HEM	CAA-CBA-CGA	-4.87	100.72	113.83
3	D	1002	HEM	C4B-CHC-C1C	3.51	127.20	122.56
3	A	1002	HEM	CAA-CBA-CGA	-3.47	104.48	113.83
3	B	1002	HEM	C4C-CHD-C1D	3.37	127.00	122.56
3	B	1002	HEM	CAA-CBA-CGA	-3.34	104.84	113.83
3	A	1002	HEM	C4B-CHC-C1C	3.32	126.94	122.56
3	D	1002	HEM	CAA-CBA-CGA	-2.88	106.07	113.83
3	C	1002	HEM	CBA-CAA-C2A	2.71	117.10	112.54
5	B	1007	GOL	C3-C2-C1	-2.66	102.05	111.80
3	A	1002	HEM	CMC-C2C-C3C	2.61	129.90	124.68
5	B	1006	GOL	C3-C2-C1	-2.57	102.35	111.80
3	C	1002	HEM	C1B-NB-C4B	2.57	108.25	105.21
3	B	1002	HEM	CMC-C2C-C3C	2.53	129.73	124.68

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	1007	GOL	C3-C2-C1	-2.49	102.68	111.80
3	C	1002	HEM	CMC-C2C-C3C	2.45	129.58	124.68
3	C	1002	HEM	C4B-CHC-C1C	2.42	125.75	122.56
3	C	1002	HEM	C3B-C2B-C1B	2.35	108.18	106.41
5	A	1006	GOL	C3-C2-C1	-2.28	103.44	111.80
3	A	1002	HEM	O2D-CGD-CBD	2.26	121.13	114.00
3	A	1002	HEM	C4A-C3A-C2A	2.20	108.53	107.00
3	A	1002	HEM	C1B-NB-C4B	2.20	107.81	105.21
5	B	1009	GOL	C3-C2-C1	-2.18	103.82	111.80
5	B	1008	GOL	C3-C2-C1	-2.17	103.83	111.80
3	B	1002	HEM	C3B-C2B-C1B	2.15	108.03	106.41
3	B	1002	HEM	C4A-C3A-C2A	2.10	108.46	107.00
3	D	1002	HEM	C4C-CHD-C1D	2.10	125.33	122.56
3	B	1002	HEM	CBA-CAA-C2A	2.10	116.06	112.54
5	B	1005	GOL	C3-C2-C1	-2.06	104.26	111.80
5	C	1004	GOL	C3-C2-C1	-2.03	104.34	111.80
3	B	1002	HEM	C4D-ND-C1D	2.01	107.59	105.21

There are no chirality outliers.

All (27) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1006	GOL	O1-C1-C2-C3
5	A	1007	GOL	O1-C1-C2-C3
5	B	1006	GOL	O1-C1-C2-C3
5	B	1008	GOL	O1-C1-C2-C3
5	D	1007	GOL	O1-C1-C2-O2
5	D	1007	GOL	O1-C1-C2-C3
5	D	1009	GOL	C1-C2-C3-O3
5	A	1004	GOL	C1-C2-C3-O3
5	B	1009	GOL	O1-C1-C2-C3
5	D	1006	GOL	C1-C2-C3-O3
5	D	1008	GOL	O1-C1-C2-C3
5	A	1007	GOL	O1-C1-C2-O2
5	B	1006	GOL	O1-C1-C2-O2
5	D	1009	GOL	O2-C2-C3-O3
5	A	1006	GOL	O1-C1-C2-O2
5	B	1008	GOL	O1-C1-C2-O2
5	B	1009	GOL	O1-C1-C2-O2
3	A	1002	HEM	CAA-CBA-CGA-O1A
3	A	1002	HEM	CAA-CBA-CGA-O2A
3	C	1002	HEM	CAA-CBA-CGA-O2A

*Continued on next page...*

*Continued from previous page...*

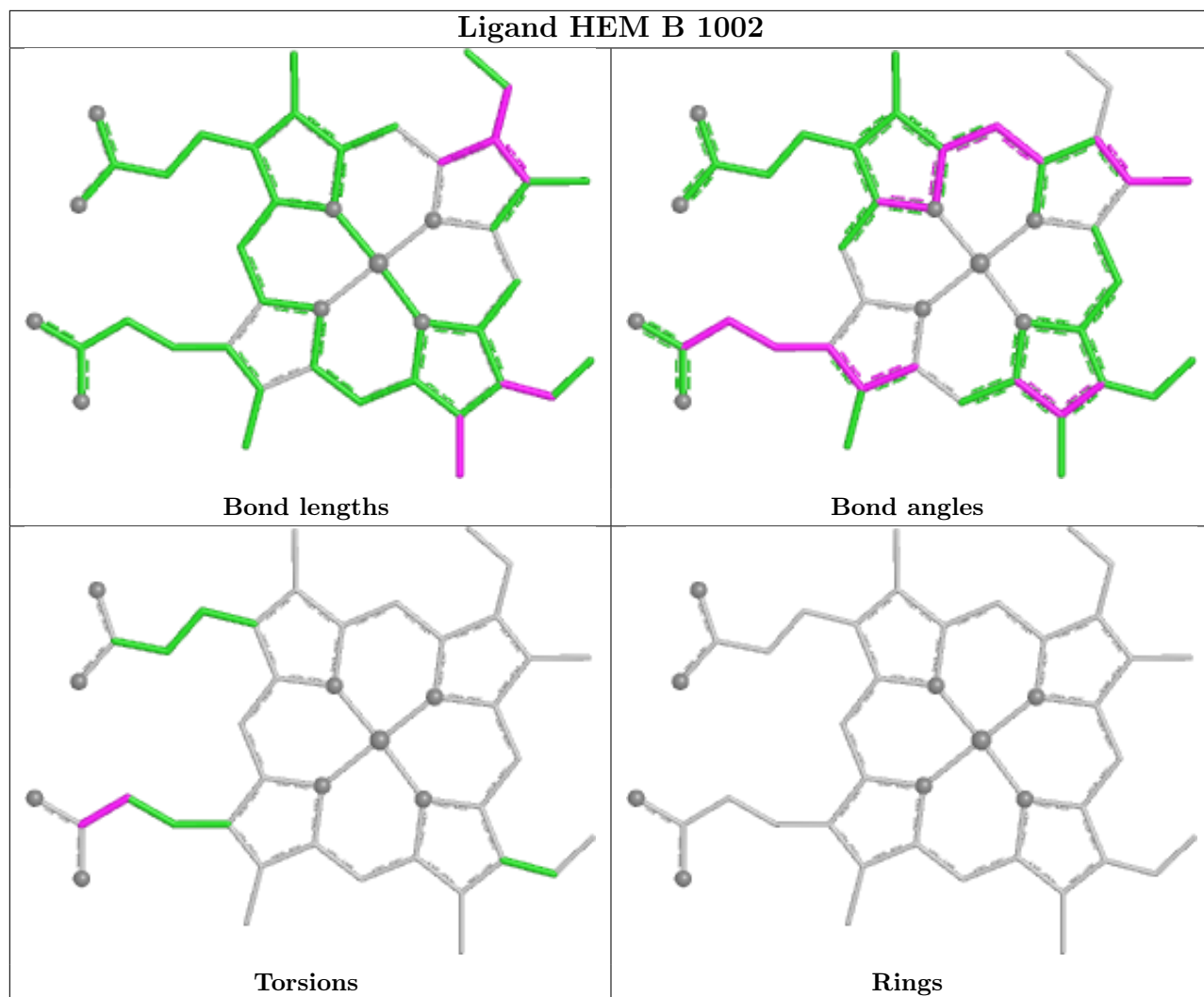
Mol	Chain	Res	Type	Atoms
3	C	1002	HEM	CAA-CBA-CGA-O1A
3	B	1002	HEM	CAA-CBA-CGA-O2A
3	D	1002	HEM	CAA-CBA-CGA-O2A
3	B	1002	HEM	CAA-CBA-CGA-O1A
3	D	1002	HEM	CAA-CBA-CGA-O1A
5	D	1004	GOL	C1-C2-C3-O3
3	C	1002	HEM	CAD-CBD-CGD-O2D

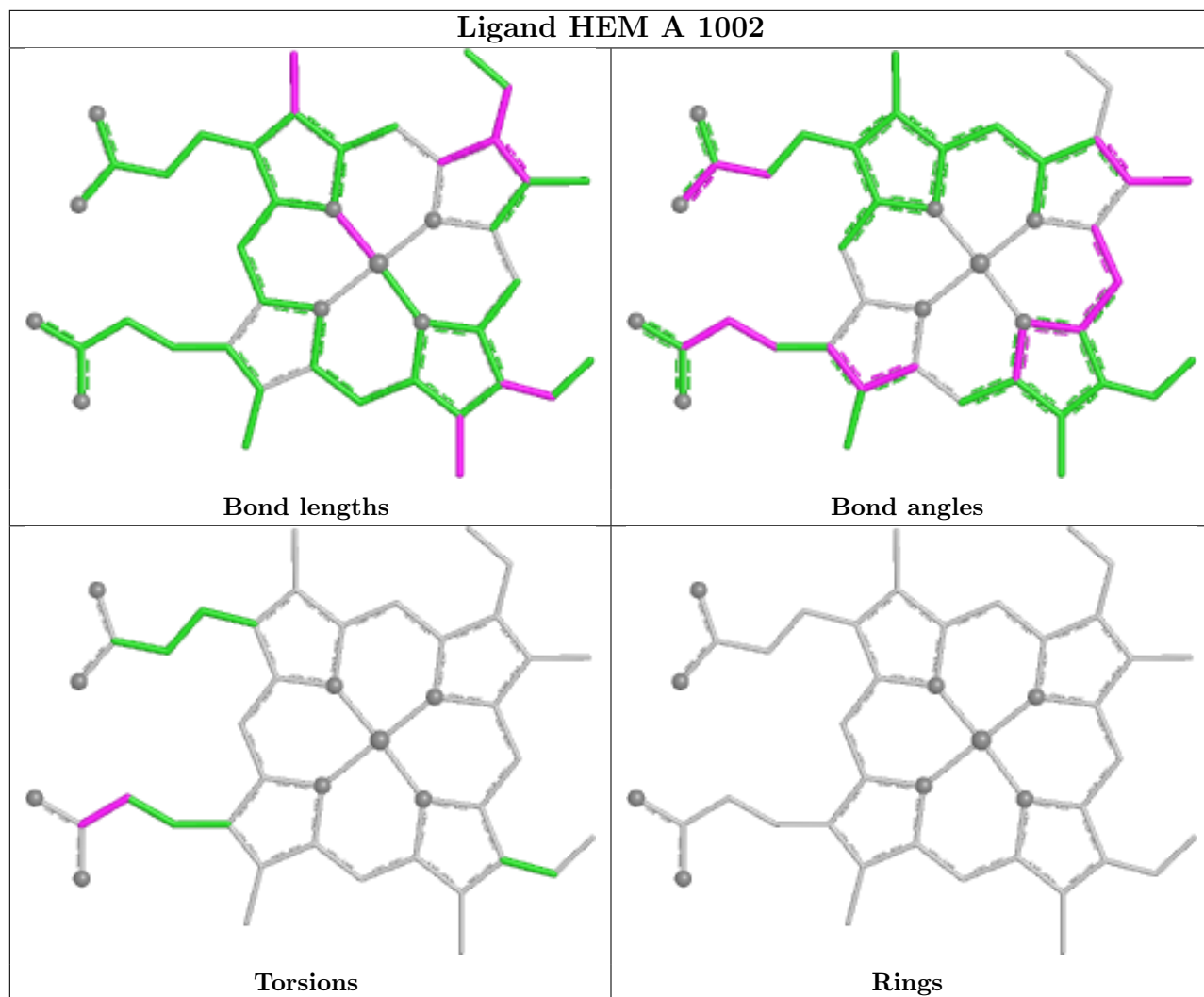
There are no ring outliers.

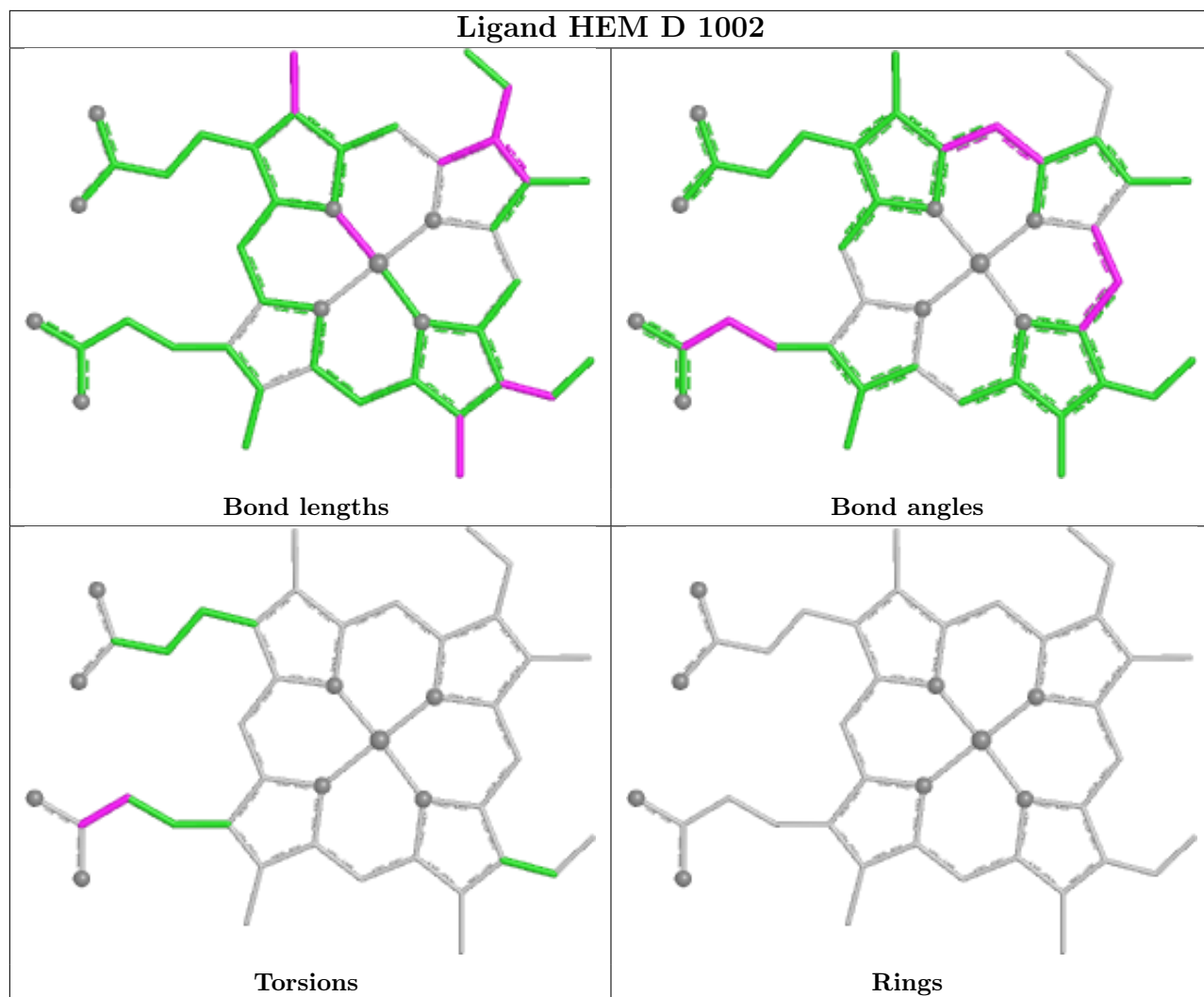
8 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1007	GOL	2	0
5	A	1004	GOL	1	0
5	D	1007	GOL	2	0
5	B	1006	GOL	1	0
4	D	1003	ACT	1	0
5	D	1009	GOL	2	0
5	B	1008	GOL	2	0
5	B	1009	GOL	1	0

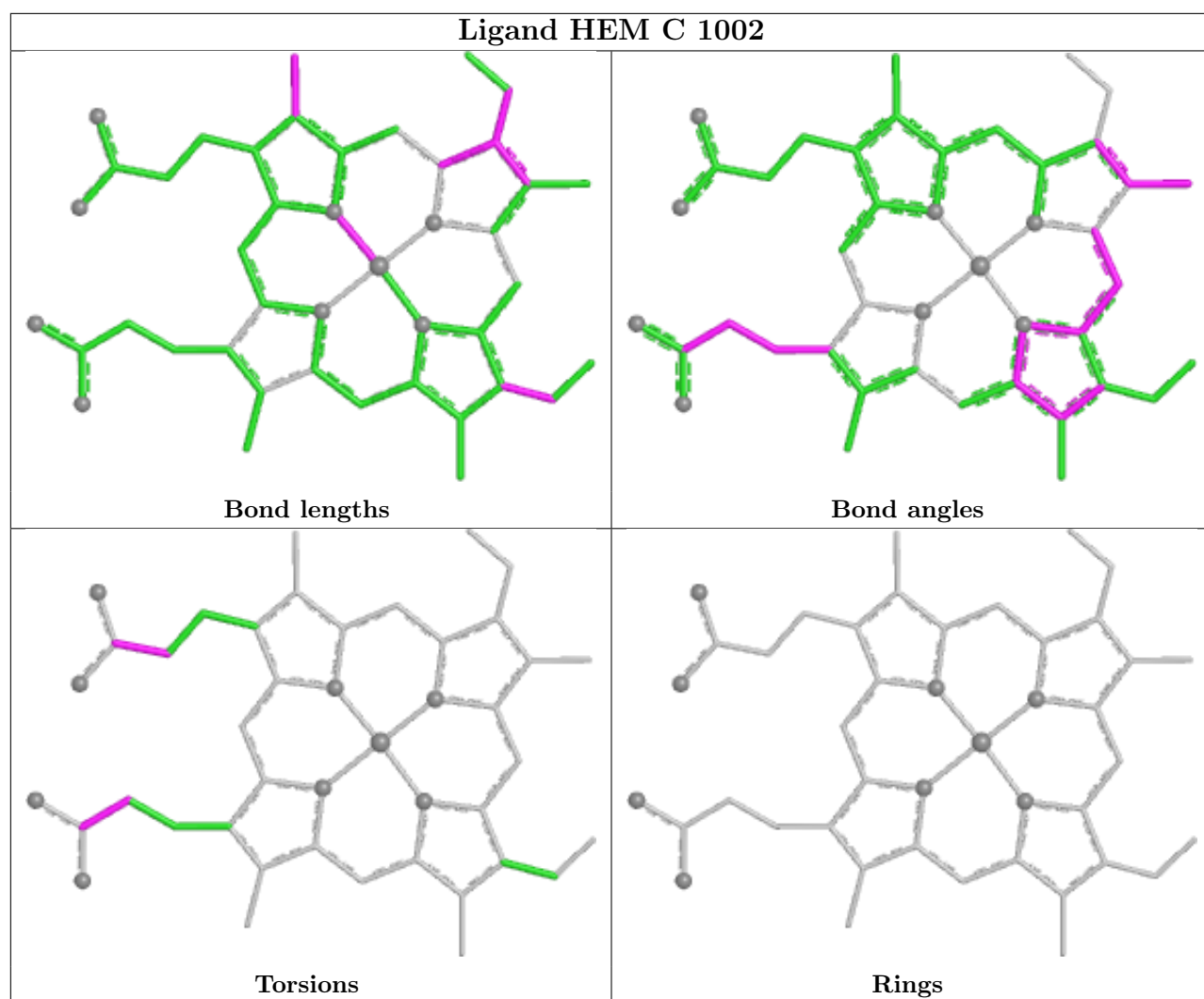
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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	717/741 (96%)	-0.01	14 (1%) 64 63	14, 20, 34, 51	0
1	B	717/741 (96%)	-0.01	5 (0%) 84 84	14, 20, 34, 51	0
1	C	717/741 (96%)	-0.11	4 (0%) 85 85	13, 19, 33, 49	1 (0%)
1	D	717/741 (96%)	-0.10	7 (0%) 79 78	13, 19, 32, 46	0
All	All	2868/2964 (96%)	-0.06	30 (1%) 79 78	13, 20, 33, 51	1 (0%)

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	533	PRO	4.7
1	A	531	ALA	4.3
1	B	533	PRO	4.0
1	A	534	GLY	3.1
1	D	533	PRO	3.0
1	A	560	GLY	2.9
1	D	512	GLY	2.9
1	A	533	PRO	2.9
1	B	532	ALA	2.8
1	B	59	GLY	2.7
1	A	511	ASP	2.7
1	C	683	GLY	2.7
1	B	51	ASN	2.7
1	A	683	GLY	2.6
1	A	510	PRO	2.5
1	A	561	HIS	2.5
1	A	532	ALA	2.5
1	C	511	ASP	2.4
1	A	59	GLY	2.3
1	B	683	GLY	2.3
1	D	562	ASN	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	561	HIS	2.3
1	D	24	GLY	2.3
1	A	512	GLY	2.2
1	A	24	GLY	2.2
1	D	563	ILE	2.2
1	C	60	ALA	2.1
1	A	51	ASN	2.1
1	A	369	GLY	2.0
1	D	511	ASP	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( $\text{\AA}^2$ )	Q<0.9
4	ACT	A	1003	4/4	0.84	0.15	31,40,41,43	0
4	ACT	B	1004	4/4	0.84	0.15	31,35,36,36	0
5	GOL	D	1009	6/6	0.84	0.13	26,33,38,42	0
5	GOL	B	1009	6/6	0.85	0.18	21,28,30,38	0
5	GOL	B	1007	6/6	0.88	0.15	20,26,28,37	0
5	GOL	B	1008	6/6	0.91	0.10	29,29,32,34	0
5	GOL	B	1006	6/6	0.91	0.12	17,22,26,29	0
5	GOL	A	1007	6/6	0.91	0.10	20,24,24,26	0
5	GOL	D	1007	6/6	0.92	0.14	22,26,28,29	0
4	ACT	D	1003	4/4	0.92	0.15	24,25,25,28	0
5	GOL	A	1006	6/6	0.93	0.08	32,32,34,39	0
5	GOL	D	1006	6/6	0.93	0.11	18,24,26,28	0
4	ACT	B	1003	4/4	0.93	0.12	19,23,24,24	0
5	GOL	B	1005	6/6	0.93	0.09	21,24,27,28	0

*Continued on next page...*

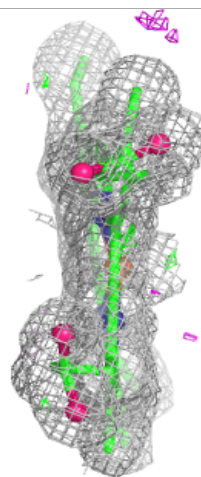
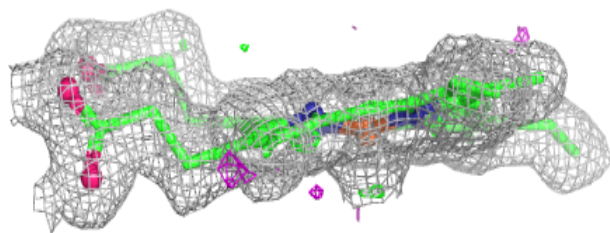
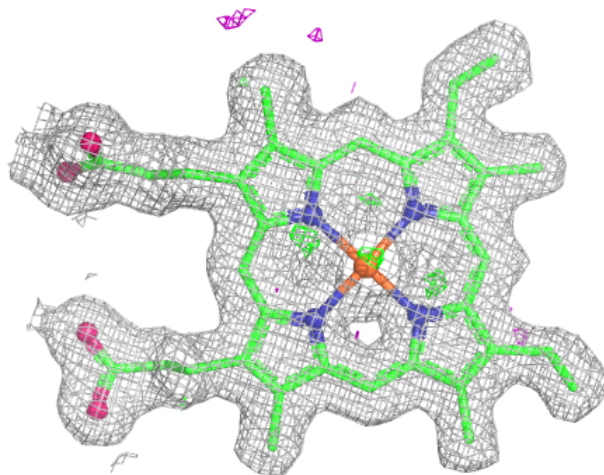
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	GOL	D	1005	6/6	0.94	0.11	19,22,23,25	0
5	GOL	A	1005	6/6	0.94	0.11	20,23,27,30	0
5	GOL	D	1008	6/6	0.95	0.11	18,25,27,30	0
5	GOL	C	1004	6/6	0.95	0.11	18,23,25,27	0
5	GOL	A	1004	6/6	0.96	0.07	17,19,20,22	0
5	GOL	D	1004	6/6	0.97	0.07	18,19,20,21	0
5	GOL	C	1003	6/6	0.97	0.05	16,17,19,20	0
3	HEM	A	1002	43/43	0.98	0.06	11,15,17,21	0
3	HEM	B	1002	43/43	0.98	0.06	12,18,21,22	0
3	HEM	C	1002	43/43	0.98	0.06	11,15,17,18	0
3	HEM	D	1002	43/43	0.98	0.06	11,14,17,19	0
2	NA	C	1001	1/1	0.98	0.04	18,18,18,18	0
2	NA	A	1001	1/1	0.99	0.03	18,18,18,18	0
2	NA	D	1001	1/1	0.99	0.03	18,18,18,18	0
2	NA	B	1001	1/1	0.99	0.03	20,20,20,20	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.

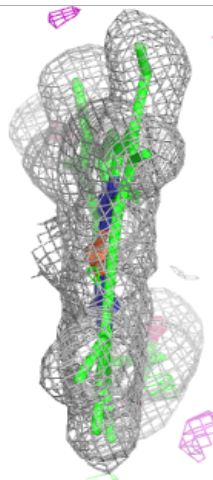
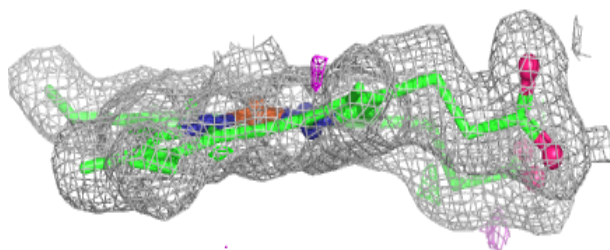
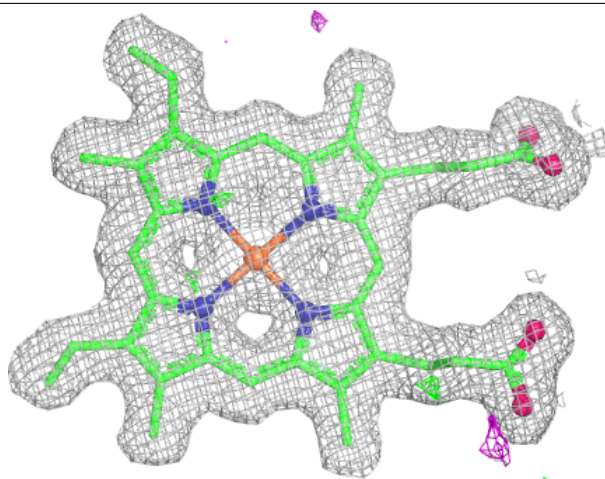
**Electron density around HEM A 1002:**

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



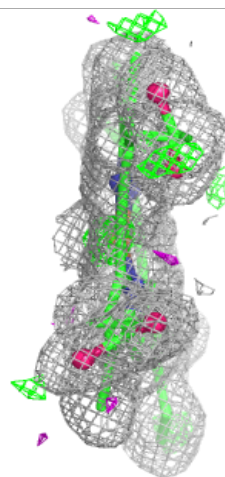
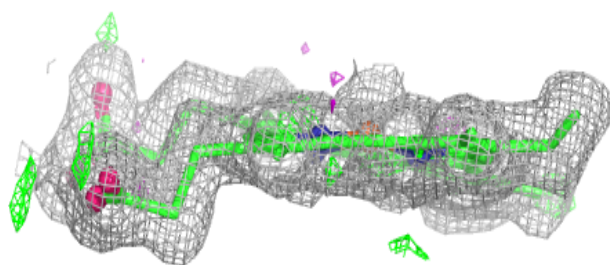
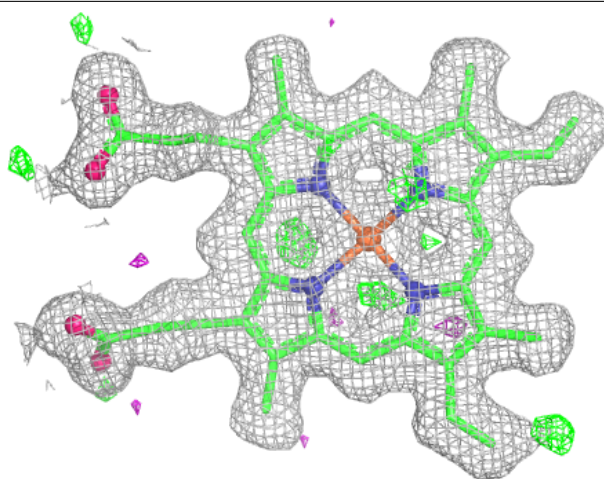
**Electron density around HEM B 1002:**

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

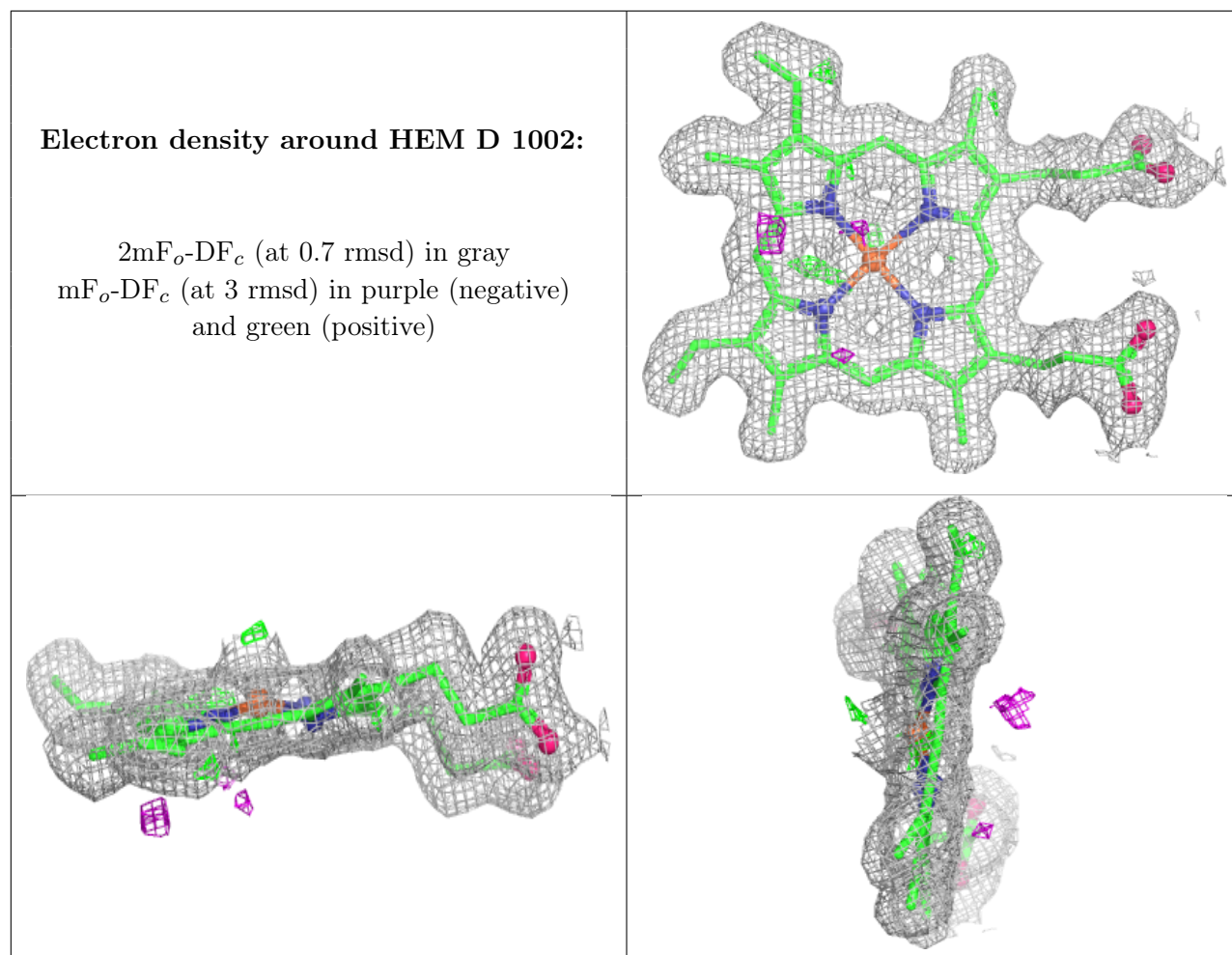


**Electron density around HEM C 1002:**

$2mF_o-DF_c$  (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.