Full wwPDB X-ray Structure Validation Report ⓘ

Jul 27, 2023 – 10:10 AM EDT

PDB ID	:	8DWR
Title	:	Crystal structure of the L333V variant of catalase-peroxidase from Mycobacterium tuberculosis
Authors	:	Diaz-Vilchis, A.; Uribe-Vazquez, B.; Avila-Linares, A.; Rudino-Pinera, E.; Soberon, X.
Deposited on	:	2022-08-01
Resolution	:	2.10 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*
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:

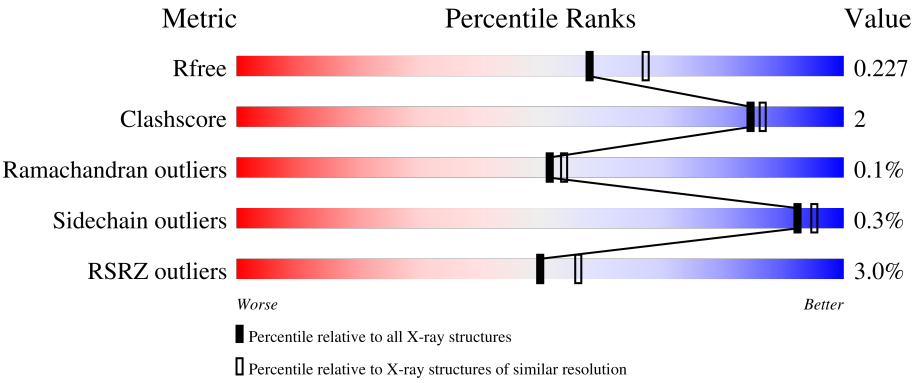
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.34
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 >=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 <=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	746	2% 91% 5% .
1	B	746	2% 92% . .
1	C	746	3% 90% 6% .
1	D	746	4% 90% 6% .

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 24593 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
1	A	717	Total	C	N	O	S	0	0	0
			5531	3514	953	1045	19			
1	B	717	Total	C	N	O	S	0	0	0
			5531	3514	953	1045	19			
1	C	717	Total	C	N	O	S	0	0	0
			5531	3514	953	1045	19			
1	D	717	Total	C	N	O	S	0	0	0
			5531	3514	953	1045	19			

There are 32 discrepancies between the modelled and reference sequences:

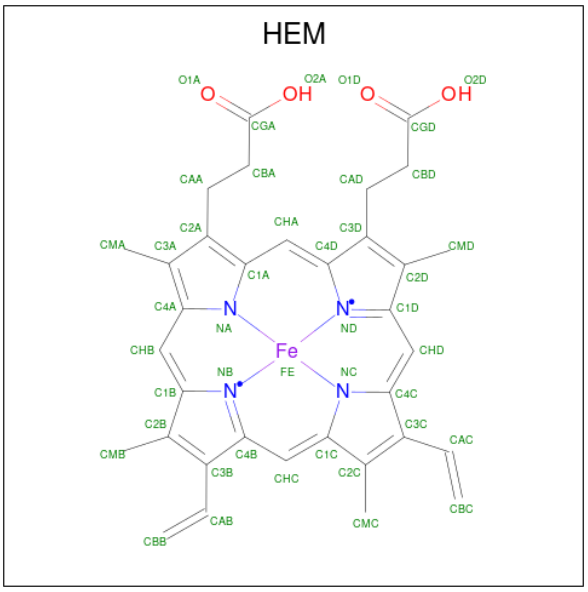
Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	MET	-	expression tag	UNP P9WIE5
A	-4	HIS	-	expression tag	UNP P9WIE5
A	-3	HIS	-	expression tag	UNP P9WIE5
A	-2	HIS	-	expression tag	UNP P9WIE5
A	-1	HIS	-	expression tag	UNP P9WIE5
A	0	HIS	-	expression tag	UNP P9WIE5
A	1	HIS	-	expression tag	UNP P9WIE5
A	333	VAL	LEU	engineered mutation	UNP P9WIE5
B	-5	MET	-	expression tag	UNP P9WIE5
B	-4	HIS	-	expression tag	UNP P9WIE5
B	-3	HIS	-	expression tag	UNP P9WIE5
B	-2	HIS	-	expression tag	UNP P9WIE5
B	-1	HIS	-	expression tag	UNP P9WIE5
B	0	HIS	-	expression tag	UNP P9WIE5
B	1	HIS	-	expression tag	UNP P9WIE5
B	333	VAL	LEU	engineered mutation	UNP P9WIE5
C	-5	MET	-	expression tag	UNP P9WIE5
C	-4	HIS	-	expression tag	UNP P9WIE5
C	-3	HIS	-	expression tag	UNP P9WIE5
C	-2	HIS	-	expression tag	UNP P9WIE5
C	-1	HIS	-	expression tag	UNP P9WIE5

Continued on next page...

Continued from previous page...

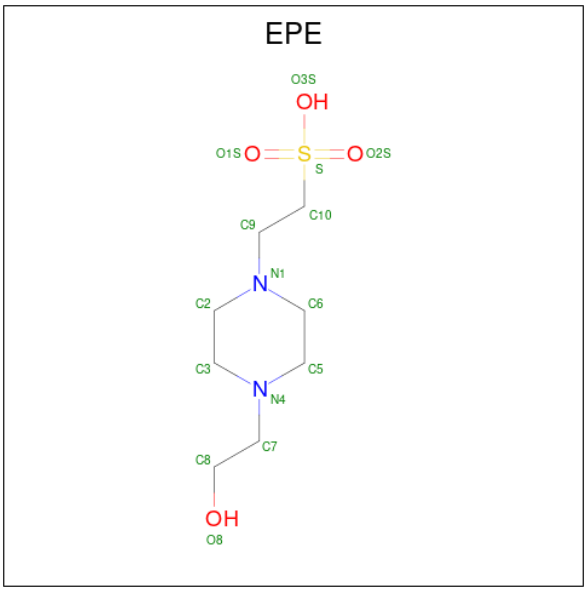
Chain	Residue	Modelled	Actual	Comment	Reference
C	0	HIS	-	expression tag	UNP P9WIE5
C	1	HIS	-	expression tag	UNP P9WIE5
C	333	VAL	LEU	engineered mutation	UNP P9WIE5
D	-5	MET	-	expression tag	UNP P9WIE5
D	-4	HIS	-	expression tag	UNP P9WIE5
D	-3	HIS	-	expression tag	UNP P9WIE5
D	-2	HIS	-	expression tag	UNP P9WIE5
D	-1	HIS	-	expression tag	UNP P9WIE5
D	0	HIS	-	expression tag	UNP P9WIE5
D	1	HIS	-	expression tag	UNP P9WIE5
D	333	VAL	LEU	engineered mutation	UNP P9WIE5

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

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).

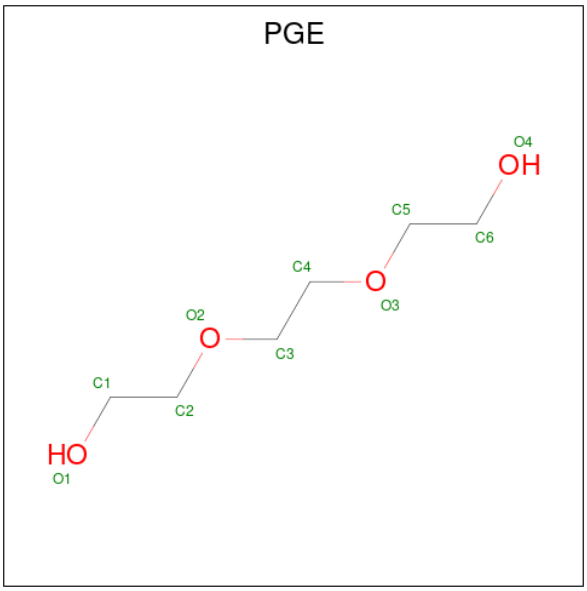


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O S 15 8 2 4 1	0	0
3	B	1	Total C N O S 15 8 2 4 1	0	0
3	C	1	Total C N O S 15 8 2 4 1	0	0
3	C	1	Total C N O S 15 8 2 4 1	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 5 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 10 6 4	0	0
5	B	1	Total C O 10 6 4	0	0
5	C	1	Total C O 10 6 4	0	0
5	D	1	Total C O 10 6 4	0	0

- Molecule 6 is water.

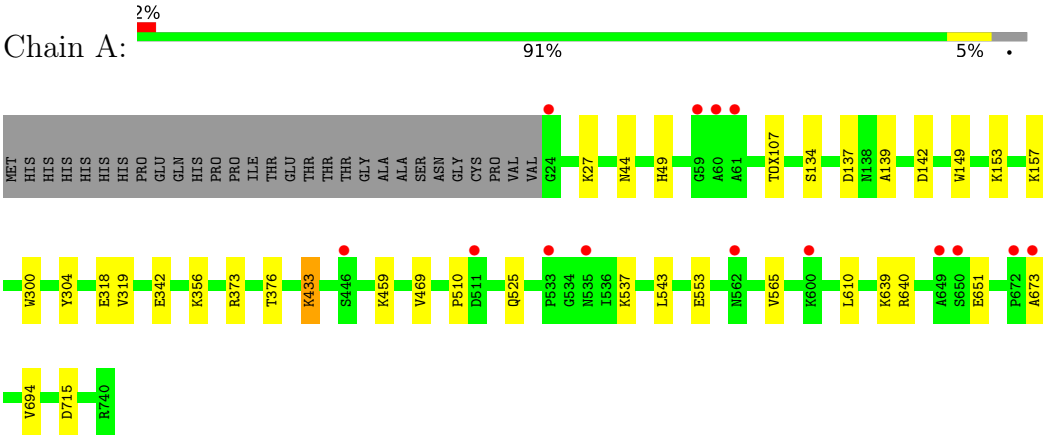
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	608	Total O 608 608	0	0
6	B	595	Total O 595 595	0	0
6	C	503	Total O 503 503	0	0
6	D	487	Total O 487 487	0	0

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.

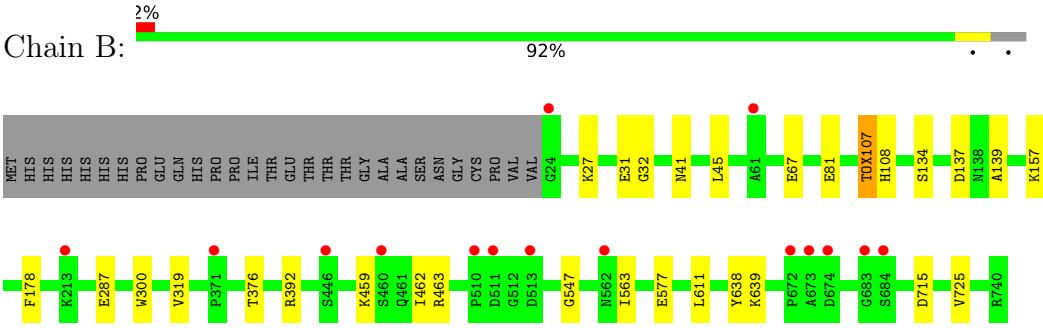
• Molecule 1: Catalase-peroxidase

Chain A:



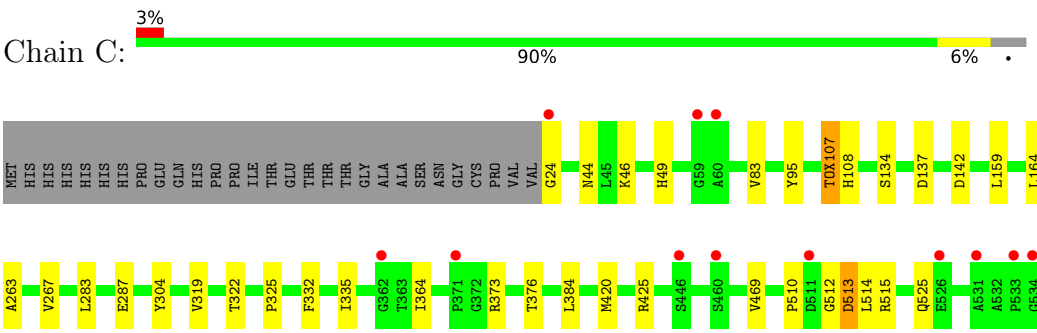
• Molecule 1: Catalase-peroxidase

Chain B:

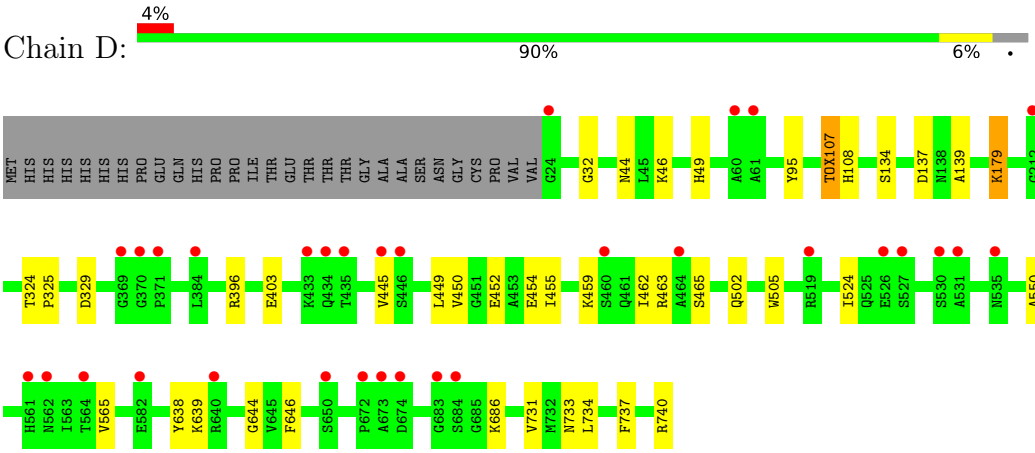


• Molecule 1: Catalase-peroxidase

Chain C:



• Molecule 1: Catalase-peroxidase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	82.67Å 150.82Å 127.04Å 90.00° 90.45° 90.00°	Depositor
Resolution (Å)	19.99 – 2.10 19.99 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.0 (19.99-2.10) 99.0 (19.99-2.10)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.00 (at 2.09Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.181 , 0.227 0.181 , 0.227	Depositor DCC
R_{free} test set	2007 reflections (1.12%)	wwPDB-VP
Wilson B-factor (Å^2)	23.0	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 57.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.022 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	24593	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.25% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: TOX, HEM, NA, PGE, EPE

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	0.25	0/5663	0.42	0/7705
1	B	0.25	0/5663	0.43	0/7705
1	C	0.25	0/5663	0.42	0/7705
1	D	0.26	0/5663	0.43	0/7705
All	All	0.25	0/22652	0.43	0/30820

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	5531	0	5352	26	0
1	B	5531	0	5352	22	0
1	C	5531	0	5352	26	0
1	D	5531	0	5352	30	0
2	A	43	0	30	0	0
2	B	43	0	30	0	0
2	C	43	0	30	0	0
2	D	43	0	30	0	0
3	A	15	0	17	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	15	0	17	1	0
3	C	30	0	34	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
5	A	10	0	14	1	0
5	B	10	0	14	2	0
5	C	10	0	14	1	0
5	D	10	0	14	0	0
6	A	608	0	0	4	0
6	B	595	0	0	3	0
6	C	503	0	0	1	0
6	D	487	0	0	3	0
All	All	24593	0	21652	102	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 (102) 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:510:PRO:HB2	1:A:639:LYS:HZ3	1.45	0.81
1:D:733:ASN:O	1:D:740:ARG:NH1	2.20	0.73
1:D:553:GLU:HG3	1:D:565:VAL:HG23	1.75	0.68
1:D:450:VAL:HG12	1:D:455:ILE:HG13	1.77	0.65
1:A:537:LYS:NZ	6:A:905:HOH:O	2.30	0.63
1:A:510:PRO:HB2	1:A:639:LYS:NZ	2.15	0.61
1:C:513:ASP:O	1:C:515:ARG:N	2.34	0.61
1:B:392:ARG:NH2	6:B:907:HOH:O	2.34	0.60
1:C:24:GLY:N	6:C:910:HOH:O	2.35	0.58
1:D:450:VAL:HG13	1:D:454:GLU:HB2	1.85	0.58
1:B:27:LYS:HB3	1:B:31:GLU:HG3	1.85	0.58
1:D:686:LYS:NZ	6:D:911:HOH:O	2.34	0.57
1:D:324:THR:OG1	1:D:329:ASP:OD2	2.21	0.55
1:A:510:PRO:CB	1:A:639:LYS:HZ3	2.19	0.55
6:B:905:HOH:O	1:C:213:LYS:NZ	2.38	0.55
1:D:463:ARG:NH2	6:D:912:HOH:O	2.37	0.55
1:C:469:VAL:HB	1:C:651:GLU:HG2	1.90	0.54
1:D:459:LYS:HG2	1:D:550:ALA:HB2	1.90	0.54
1:D:396:ARG:NH1	1:D:403:GLU:OE2	2.35	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:462:ILE:O	1:D:465:SER:OG	2.20	0.53
1:D:445:VAL:HG11	1:D:449:LEU:HD21	1.91	0.53
1:A:610:LEU:HD22	1:A:694:VAL:HG13	1.92	0.52
1:C:44:ASN:HD22	1:D:32:GLY:HA2	1.73	0.52
1:C:263:ALA:O	1:C:267:VAL:HG13	2.09	0.52
1:A:553:GLU:HG3	1:A:565:VAL:HG23	1.92	0.52
1:C:610:LEU:HD22	1:C:694:VAL:HG13	1.91	0.52
1:A:44:ASN:HD22	1:B:32:GLY:HA2	1.75	0.51
1:A:342:GLU:HB3	1:A:356:LYS:HG2	1.93	0.51
1:C:364:ILE:HB	1:C:373:ARG:HB2	1.92	0.51
1:C:510:PRO:HB3	1:C:638:TYR:CE1	2.46	0.51
1:D:731:VAL:HA	1:D:734:LEU:HG	1.94	0.50
1:D:324:THR:O	6:D:901:HOH:O	2.20	0.50
1:C:83:VAL:HG23	5:C:805:PGE:H12	1.93	0.50
1:A:433:LYS:HD2	1:A:433:LYS:H	1.77	0.50
1:D:46:LYS:HA	1:D:49:HIS:CE1	2.46	0.50
1:D:452:GLU:H	1:D:452:GLU:CD	2.15	0.49
1:C:600:LYS:NZ	1:C:601:GLY:H	2.10	0.49
1:D:107:TOX:CD1	1:D:108:HIS:HD2	2.25	0.49
1:C:267:VAL:HB	1:C:332:PHE:CD1	2.47	0.49
1:C:319:VAL:HG22	1:C:376:THR:HB	1.94	0.49
1:D:450:VAL:CG1	1:D:455:ILE:HG13	2.43	0.49
1:A:318:GLU:OE2	1:A:373:ARG:NH2	2.39	0.49
1:C:267:VAL:HG12	1:C:384:LEU:HD13	1.95	0.48
1:A:459:LYS:NZ	1:A:553:GLU:OE2	2.42	0.48
1:B:67:GLU:HB3	1:B:158:LYS:HA	1.94	0.48
1:D:450:VAL:HG13	1:D:454:GLU:CB	2.44	0.47
1:D:737:PHE:HA	1:D:740:ARG:HB3	1.96	0.47
1:B:462:ILE:HD13	1:B:547:GLY:HA2	1.95	0.47
1:B:638:TYR:CZ	1:B:639:LYS:HE2	2.50	0.47
1:C:134:SER:HB3	1:C:287:GLU:HG3	1.96	0.47
1:A:640:ARG:NH1	6:A:916:HOH:O	2.37	0.47
1:B:134:SER:HB3	1:B:287:GLU:HG3	1.97	0.47
1:A:525:GLN:HB2	1:A:543:LEU:HD12	1.97	0.46
1:B:107:TOX:CD1	1:B:108:HIS:HD2	2.29	0.46
1:A:134:SER:HB3	1:A:287:GLU:HG3	1.98	0.45
1:D:179:LYS:HB3	1:D:179:LYS:HE3	1.62	0.45
1:A:673:ALA:O	1:A:675:ASP:N	2.50	0.45
1:B:45:LEU:HD23	1:B:611:LEU:HD21	1.99	0.45
1:C:525:GLN:HB2	1:C:543:LEU:HD12	1.99	0.45
1:C:159:LEU:HD21	1:C:164:LEU:HD13	2.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:139:ALA:HA	1:A:300:TRP:CZ3	2.53	0.44
1:B:41:ASN:HB2	3:B:801:EPE:H91	1.98	0.44
1:C:555:ALA:HB1	1:C:651:GLU:O	2.17	0.44
1:A:27:LYS:NZ	1:B:41:ASN:OD1	2.51	0.44
1:D:139:ALA:HA	1:D:300:TRP:CZ3	2.53	0.43
1:A:433:LYS:NZ	6:A:948:HOH:O	2.50	0.43
1:D:95:TYR:CZ	1:D:325:PRO:HG2	2.53	0.43
1:B:81:GLU:OE2	1:B:175:SER:HB2	2.17	0.43
1:D:502:GLN:HG2	1:D:505:TRP:CH2	2.54	0.43
1:C:95:TYR:CZ	1:C:325:PRO:HG2	2.54	0.43
1:A:149:TRP:CE2	1:A:153:LYS:HB2	2.53	0.43
1:B:563:ILE:HD13	1:B:725:VAL:HG21	2.01	0.43
1:B:638:TYR:OH	1:B:639:LYS:HE2	2.18	0.43
1:B:174:GLU:HA	1:B:178:PHE:O	2.18	0.42
1:D:44:ASN:HD21	1:D:46:LYS:CD	2.32	0.42
1:A:49:HIS:O	1:A:49:HIS:ND1	2.51	0.42
1:A:142:ASP:N	1:A:142:ASP:OD1	2.52	0.42
1:D:465:SER:HB3	1:D:524:ILE:HD11	2.01	0.42
1:D:134:SER:HB3	1:D:287:GLU:HG3	2.00	0.42
1:B:459:LYS:O	1:B:463:ARG:HG2	2.19	0.42
1:C:46:LYS:HA	1:C:49:HIS:CE1	2.55	0.41
1:C:107:TOX:CD1	1:C:108:HIS:HD2	2.32	0.41
1:B:139:ALA:HA	1:B:300:TRP:CZ3	2.55	0.41
1:B:577:GLU:H	1:B:577:GLU:CD	2.23	0.41
5:B:804:PGE:H52	6:B:1409:HOH:O	2.20	0.41
5:A:804:PGE:H52	6:A:1402:HOH:O	2.20	0.41
1:C:142:ASP:N	1:C:142:ASP:OD1	2.53	0.41
1:B:319:VAL:HG22	1:B:376:THR:HB	2.02	0.41
1:A:157:LYS:HG2	1:B:715:ASP:OD2	2.21	0.41
1:A:469:VAL:HB	1:A:651:GLU:CG	2.50	0.41
1:D:638:TYR:CE2	1:D:639:LYS:HG3	2.55	0.41
1:A:715:ASP:OD2	1:B:157:LYS:HG2	2.20	0.41
5:B:804:PGE:H52	5:B:804:PGE:H32	1.94	0.41
1:D:644:GLY:HA2	1:D:646:PHE:CZ	2.55	0.41
1:C:283:LEU:HD12	1:C:304:TYR:HB2	2.02	0.40
1:C:420:MET:O	1:C:425:ARG:HD3	2.21	0.40
1:B:139:ALA:HA	1:B:300:TRP:CH2	2.56	0.40
1:C:469:VAL:HG13	1:C:551:ALA:HB1	2.04	0.40
1:D:139:ALA:HA	1:D:300:TRP:CH2	2.56	0.40
1:C:322:THR:HA	1:C:335:ILE:HD12	2.02	0.40
1:A:283:LEU:HD12	1:A:304:TYR:HB2	2.03	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:319:VAL:HG22	1:A:376:THR:HB	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	714/746 (96%)	702 (98%)	12 (2%)	0	100 100
1	B	714/746 (96%)	698 (98%)	16 (2%)	0	100 100
1	C	714/746 (96%)	696 (98%)	15 (2%)	3 (0%)	34 32
1	D	714/746 (96%)	696 (98%)	18 (2%)	0	100 100
All	All	2856/2984 (96%)	2792 (98%)	61 (2%)	3 (0%)	51 54

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	513	ASP
1	C	512	GLY
1	C	514	LEU

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	564/589 (96%)	562 (100%)	2 (0%)	91	94
1	B	564/589 (96%)	563 (100%)	1 (0%)	93	96
1	C	564/589 (96%)	562 (100%)	2 (0%)	91	94
1	D	564/589 (96%)	562 (100%)	2 (0%)	91	94
All	All	2256/2356 (96%)	2249 (100%)	7 (0%)	92	95

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	137	ASP
1	A	433	LYS
1	B	137	ASP
1	C	137	ASP
1	C	214	ARG
1	D	137	ASP
1	D	179	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	TOX	A	107	2,1	10,17,18	2.27	1 (10%)	10,23,25	0.75	0
1	TOX	C	107	2,1	10,17,18	2.29	1 (10%)	10,23,25	0.75	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
1	TOX	D	107	2,1	10,17,18	2.22	1 (10%)	10,23,25	0.65	0
1	TOX	B	107	2,1	10,17,18	2.26	1 (10%)	10,23,25	0.76	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
1	TOX	A	107	2,1	-	2/4/8/10	0/2/2/2
1	TOX	C	107	2,1	-	2/4/8/10	0/2/2/2
1	TOX	D	107	2,1	-	2/4/8/10	0/2/2/2
1	TOX	B	107	2,1	-	2/4/8/10	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	107	TOX	CD1-NE1	-6.85	1.32	1.39
1	A	107	TOX	CD1-NE1	-6.81	1.32	1.39
1	B	107	TOX	CD1-NE1	-6.77	1.32	1.39
1	D	107	TOX	CD1-NE1	-6.66	1.32	1.39

There are no bond angle outliers.

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	107	TOX	N-CA-CB-CG
1	A	107	TOX	C-CA-CB-CG
1	B	107	TOX	N-CA-CB-CG
1	B	107	TOX	C-CA-CB-CG
1	C	107	TOX	N-CA-CB-CG
1	C	107	TOX	C-CA-CB-CG
1	D	107	TOX	N-CA-CB-CG
1	D	107	TOX	C-CA-CB-CG

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	107	TOX	1	0
1	D	107	TOX	1	0
1	B	107	TOX	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 16 ligands modelled in this entry, 4 are monoatomic - leaving 12 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	$\#\|Z\| > 2$
2	HEM	A	801	1	41,50,50	1.48	4 (9%)	45,82,82	1.41	7 (15%)
5	PGE	A	804	-	9,9,9	0.30	0	8,8,8	0.31	0
3	EPE	B	801	-	15,15,15	0.82	1 (6%)	18,20,20	1.69	5 (27%)
5	PGE	D	803	-	9,9,9	0.31	0	8,8,8	0.28	0
2	HEM	B	802	1	41,50,50	1.46	3 (7%)	45,82,82	1.41	7 (15%)
5	PGE	B	804	-	9,9,9	0.30	0	8,8,8	0.30	0
2	HEM	D	801	1	41,50,50	1.47	3 (7%)	45,82,82	1.41	8 (17%)
3	EPE	C	802	-	15,15,15	0.81	1 (6%)	18,20,20	1.86	5 (27%)
2	HEM	C	801	1	41,50,50	1.47	3 (7%)	45,82,82	1.40	7 (15%)
5	PGE	C	805	-	9,9,9	0.32	0	8,8,8	0.24	0
3	EPE	C	804	-	15,15,15	0.82	1 (6%)	18,20,20	1.81	7 (38%)
3	EPE	A	802	-	15,15,15	0.74	1 (6%)	18,20,20	1.69	3 (16%)

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	801	1	-	2/12/54/54	-
5	PGE	A	804	-	-	4/7/7/7	-
3	EPE	B	801	-	-	1/9/19/19	0/1/1/1
5	PGE	D	803	-	-	5/7/7/7	-
2	HEM	B	802	1	-	2/12/54/54	-
5	PGE	B	804	-	-	3/7/7/7	-
2	HEM	D	801	1	-	4/12/54/54	-
3	EPE	C	802	-	-	1/9/19/19	0/1/1/1
2	HEM	C	801	1	-	2/12/54/54	-
5	PGE	C	805	-	-	0/7/7/7	-
3	EPE	C	804	-	-	5/9/19/19	0/1/1/1
3	EPE	A	802	-	-	5/9/19/19	0/1/1/1

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	801	HEM	C3C-C2C	-3.93	1.34	1.40
2	D	801	HEM	C3C-C2C	-3.92	1.34	1.40
2	A	801	HEM	C3C-C2C	-3.87	1.35	1.40
2	A	801	HEM	C3C-CAC	3.80	1.55	1.47
2	B	802	HEM	C3C-CAC	3.79	1.55	1.47
2	D	801	HEM	C3C-CAC	3.78	1.55	1.47
2	B	802	HEM	C3C-C2C	-3.77	1.35	1.40
2	C	801	HEM	C3C-CAC	3.75	1.55	1.47
2	A	801	HEM	CAB-C3B	3.00	1.55	1.47
2	C	801	HEM	CAB-C3B	2.99	1.55	1.47
2	D	801	HEM	CAB-C3B	2.92	1.55	1.47
2	B	802	HEM	CAB-C3B	2.88	1.55	1.47
3	B	801	EPE	C10-S	2.70	1.81	1.77
3	C	802	EPE	C10-S	2.70	1.81	1.77
3	C	804	EPE	C10-S	2.65	1.81	1.77
3	A	802	EPE	C10-S	2.48	1.81	1.77
2	A	801	HEM	FE-NB	2.06	2.07	1.96

All (49) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	802	EPE	C5-N4-C3	4.12	118.10	108.83
3	A	802	EPE	C5-N4-C3	4.11	118.08	108.83
3	B	801	EPE	C5-N4-C3	3.87	117.54	108.83
3	C	802	EPE	C7-N4-C5	3.86	121.10	111.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	804	EPE	C5-N4-C3	3.68	117.12	108.83
3	C	804	EPE	C7-N4-C3	3.29	119.64	111.23
3	A	802	EPE	C7-N4-C3	3.23	119.51	111.23
2	D	801	HEM	C4D-ND-C1D	3.17	108.35	105.07
2	B	802	HEM	C4D-ND-C1D	3.14	108.32	105.07
3	C	804	EPE	C7-N4-C5	3.08	119.11	111.23
2	C	801	HEM	C4D-ND-C1D	3.08	108.25	105.07
3	C	802	EPE	C7-N4-C3	2.97	118.82	111.23
2	A	801	HEM	C4D-ND-C1D	2.96	108.13	105.07
3	B	801	EPE	C7-N4-C3	2.94	118.76	111.23
2	A	801	HEM	CAA-CBA-CGA	-2.83	105.84	113.76
2	A	801	HEM	C4B-CHC-C1C	2.76	126.20	122.56
3	C	802	EPE	O3S-S-C10	2.65	110.06	105.77
2	B	802	HEM	CAA-CBA-CGA	-2.65	106.33	113.76
2	C	801	HEM	CAA-CBA-CGA	-2.59	106.49	113.76
2	D	801	HEM	C4C-CHD-C1D	2.57	125.95	122.56
2	B	802	HEM	C4C-CHD-C1D	2.56	125.94	122.56
2	C	801	HEM	C1B-NB-C4B	2.55	107.71	105.07
2	D	801	HEM	C1B-NB-C4B	2.54	107.70	105.07
2	B	802	HEM	C1B-NB-C4B	2.54	107.70	105.07
2	A	801	HEM	C1B-NB-C4B	2.52	107.68	105.07
3	C	802	EPE	O2S-S-C10	2.52	109.95	106.92
3	B	801	EPE	C7-N4-C5	2.47	117.56	111.23
3	A	802	EPE	O3S-S-C10	2.47	109.76	105.77
2	C	801	HEM	C4B-CHC-C1C	2.46	125.81	122.56
2	B	802	HEM	C4B-CHC-C1C	2.43	125.76	122.56
3	B	801	EPE	O3S-S-C10	2.36	109.59	105.77
2	C	801	HEM	C3D-C4D-ND	-2.31	107.59	110.17
2	D	801	HEM	C3D-C4D-ND	-2.28	107.63	110.17
3	C	804	EPE	O1S-S-C10	2.28	109.66	106.92
2	A	801	HEM	C3D-C4D-ND	-2.28	107.63	110.17
2	C	801	HEM	C4C-CHD-C1D	2.25	125.53	122.56
2	D	801	HEM	CAA-CBA-CGA	-2.25	107.46	113.76
2	B	802	HEM	C3D-C4D-ND	-2.24	107.67	110.17
2	D	801	HEM	CMC-C2C-C3C	2.24	128.86	124.68
3	C	804	EPE	O2S-S-C10	2.24	109.61	106.92
2	D	801	HEM	C3B-C2B-C1B	2.20	108.11	106.49
3	C	804	EPE	C6-N1-C2	2.17	113.70	108.83
2	D	801	HEM	C4B-CHC-C1C	2.14	125.38	122.56
2	B	802	HEM	CMC-C2C-C3C	2.14	128.68	124.68
3	B	801	EPE	O2S-S-C10	2.13	109.48	106.92
3	C	804	EPE	O3S-S-C10	2.06	109.10	105.77

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	801	HEM	CMC-C2C-C3C	2.04	128.50	124.68
2	A	801	HEM	C4C-CHD-C1D	2.02	125.23	122.56
2	C	801	HEM	CMC-C2C-C3C	2.01	128.44	124.68

There are no chirality outliers.

All (34) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	802	EPE	C8-C7-N4-C3
3	A	802	EPE	C9-C10-S-O2S
3	A	802	EPE	C9-C10-S-O3S
3	B	801	EPE	C8-C7-N4-C3
3	C	802	EPE	C8-C7-N4-C5
3	C	804	EPE	C9-C10-S-O1S
3	C	804	EPE	C9-C10-S-O2S
3	C	804	EPE	C9-C10-S-O3S
5	B	804	PGE	O1-C1-C2-O2
5	D	803	PGE	O2-C3-C4-O3
5	D	803	PGE	O3-C5-C6-O4
5	D	803	PGE	O1-C1-C2-O2
5	A	804	PGE	O2-C3-C4-O3
5	B	804	PGE	O2-C3-C4-O3
5	A	804	PGE	O3-C5-C6-O4
3	A	802	EPE	S-C10-C9-N1
3	C	804	EPE	N4-C7-C8-O8
5	A	804	PGE	C1-C2-O2-C3
3	A	802	EPE	C9-C10-S-O1S
5	A	804	PGE	O1-C1-C2-O2
5	D	803	PGE	C1-C2-O2-C3
5	D	803	PGE	C3-C4-O3-C5
3	C	804	EPE	C10-C9-N1-C6
2	D	801	HEM	CAA-CBA-CGA-O2A
2	C	801	HEM	CAA-CBA-CGA-O2A
2	C	801	HEM	CAA-CBA-CGA-O1A
5	B	804	PGE	O3-C5-C6-O4
2	A	801	HEM	CAA-CBA-CGA-O2A
2	B	802	HEM	CAA-CBA-CGA-O2A
2	A	801	HEM	CAA-CBA-CGA-O1A
2	D	801	HEM	CAA-CBA-CGA-O1A
2	B	802	HEM	CAA-CBA-CGA-O1A
2	D	801	HEM	CAD-CBD-CGD-O2D
2	D	801	HEM	CAD-CBD-CGD-O1D

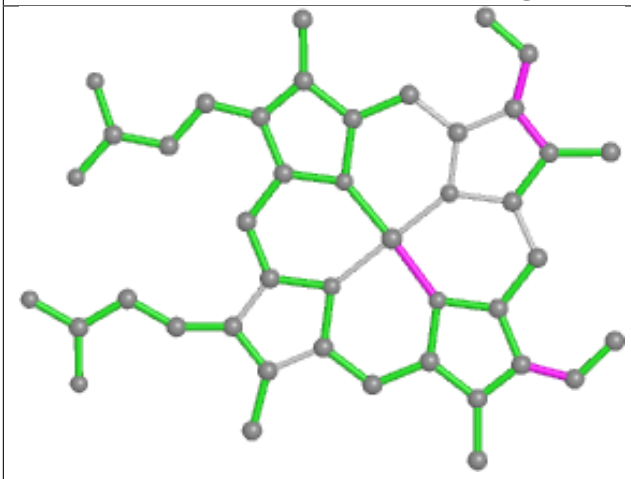
There are no ring outliers.

4 monomers are involved in 5 short contacts:

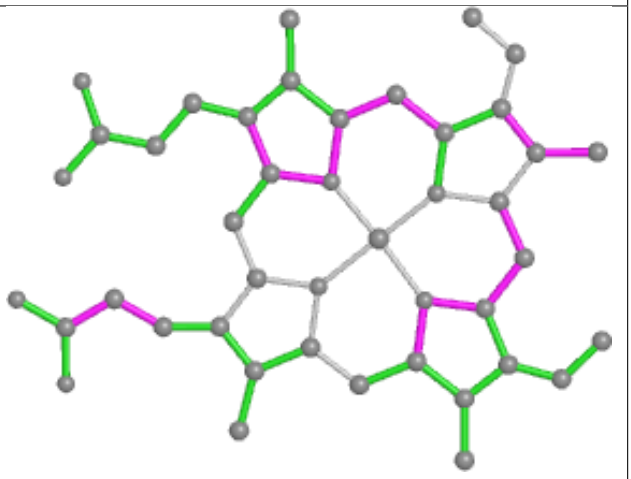
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	804	PGE	1	0
3	B	801	EPE	1	0
5	B	804	PGE	2	0
5	C	805	PGE	1	0

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 then 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.

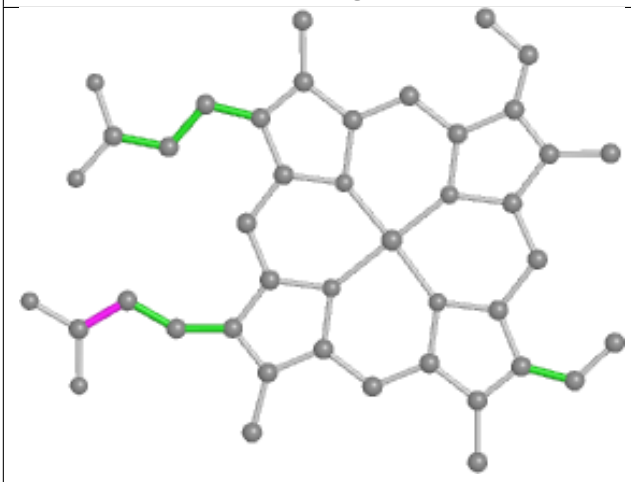
Ligand HEM A 801



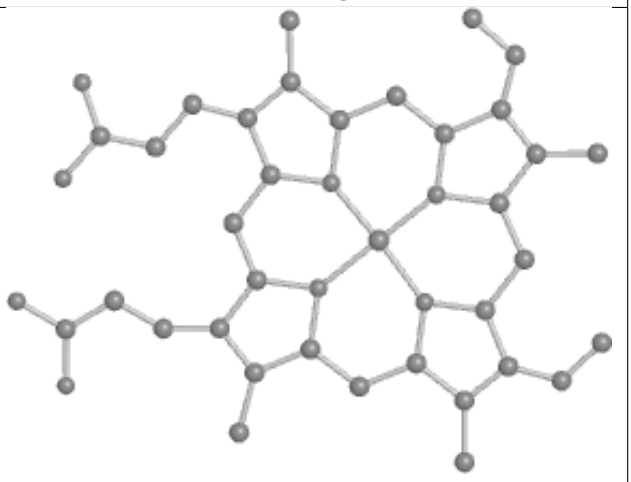
Bond lengths



Bond angles

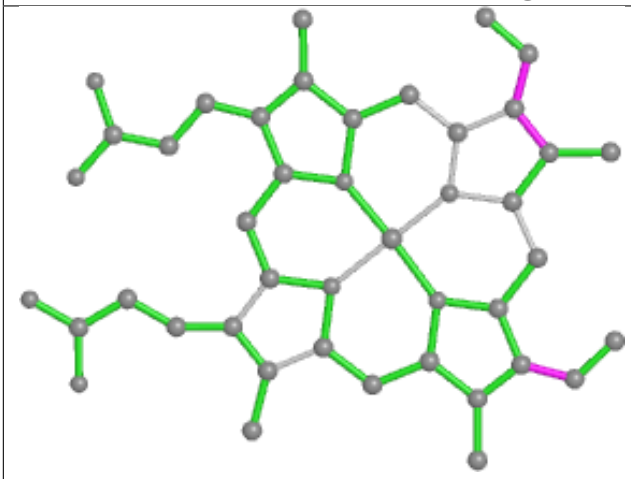


Torsions

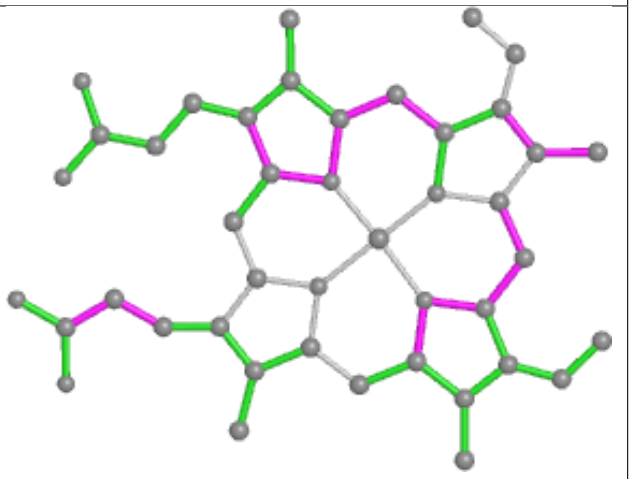


Rings

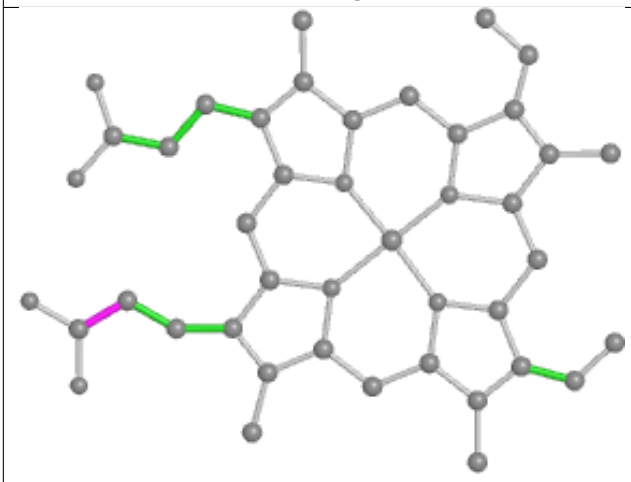
Ligand HEM B 802



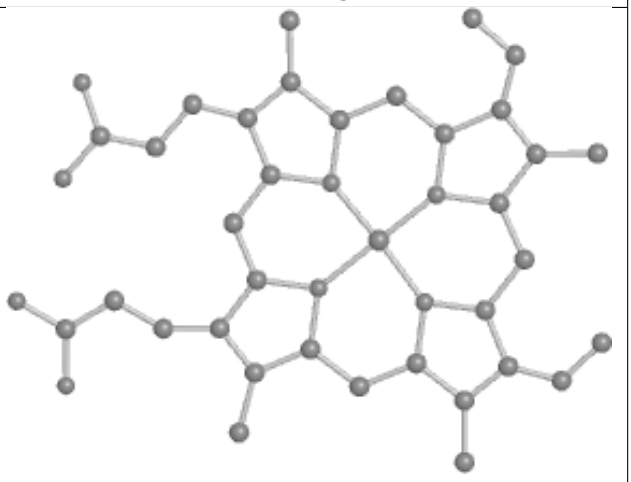
Bond lengths



Bond angles

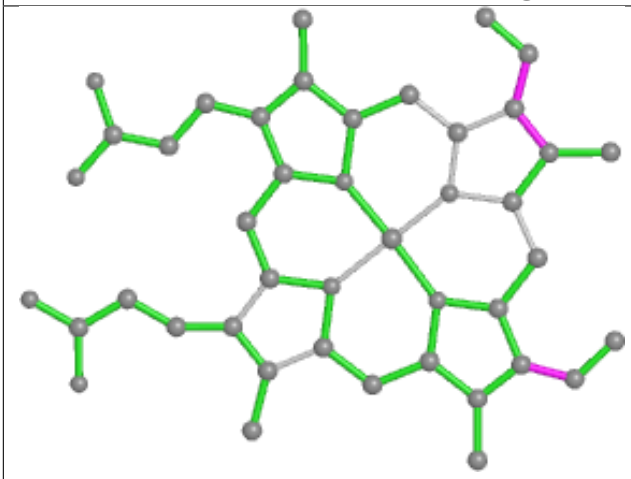


Torsions

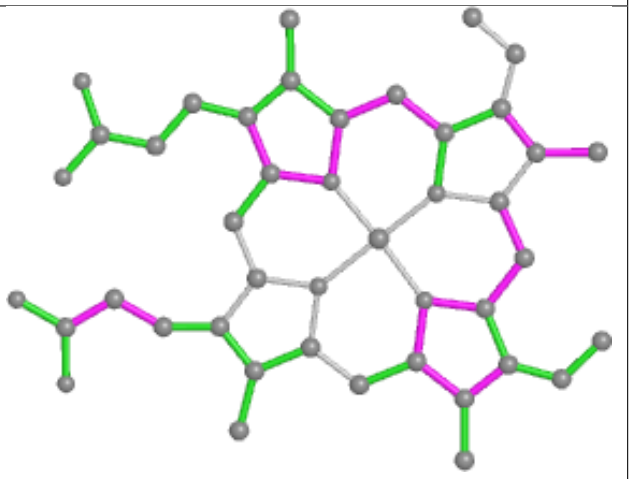


Rings

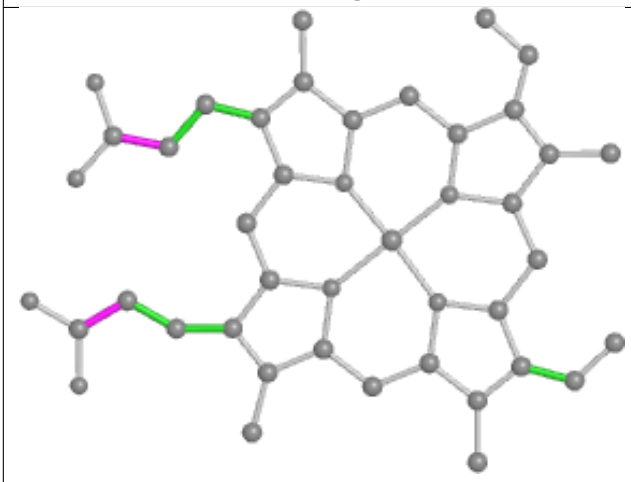
Ligand HEM D 801



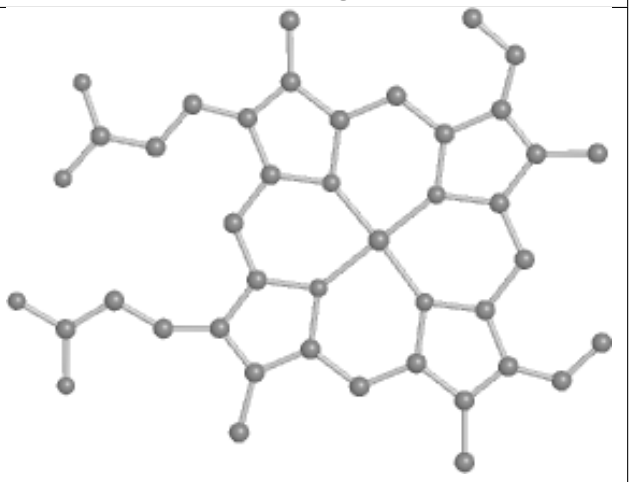
Bond lengths



Bond angles

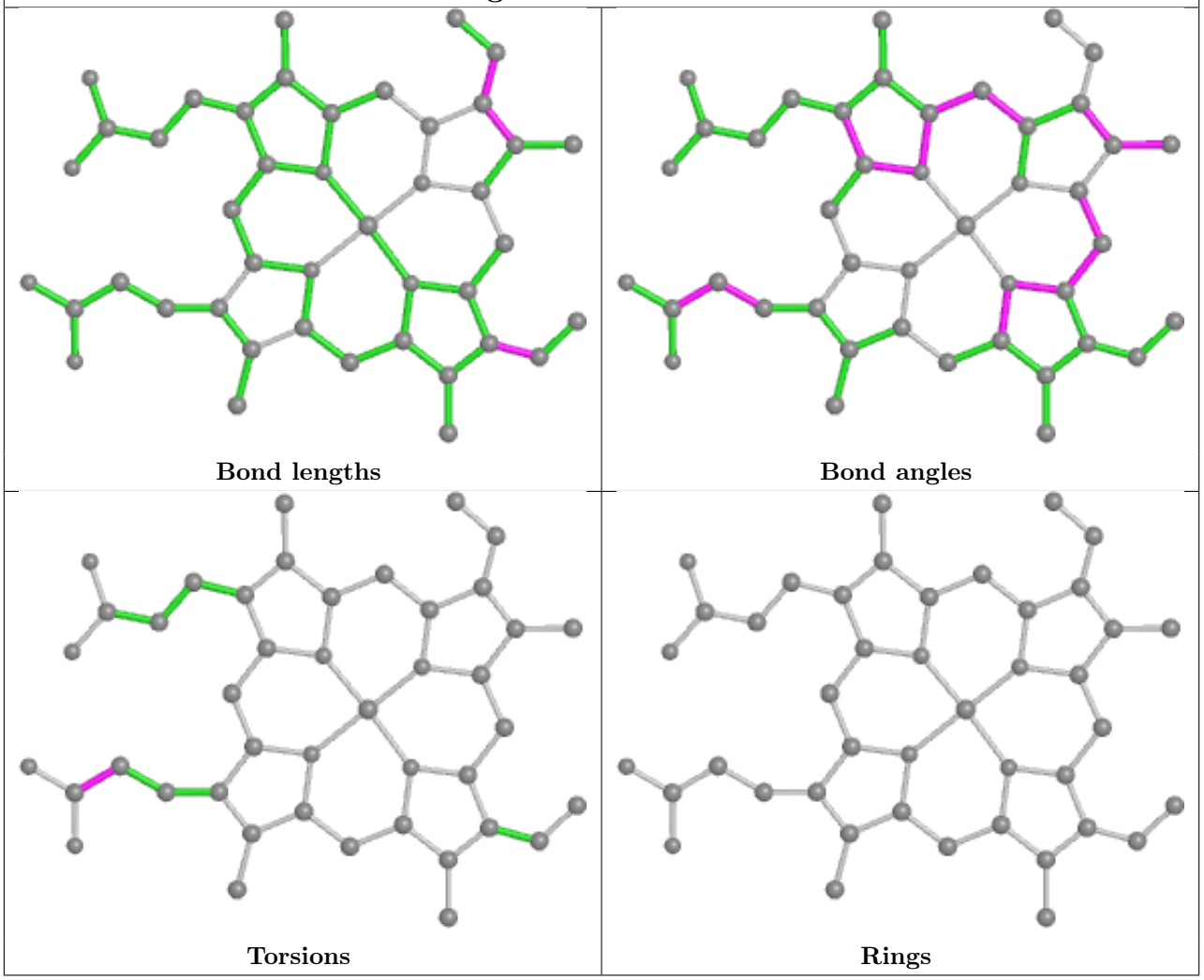


Torsions



Rings

Ligand HEM C 801



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^{th} 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^2)	Q<0.9
1	A	716/746 (95%)	-0.26	16 (2%)	62	66	12, 20, 37, 60	0
1	B	716/746 (95%)	-0.25	15 (2%)	63	68	9, 19, 35, 61	0
1	C	716/746 (95%)	-0.09	23 (3%)	47	54	13, 25, 43, 73	0
1	D	716/746 (95%)	0.08	33 (4%)	32	38	16, 27, 47, 60	0
All	All	2864/2984 (95%)	-0.13	87 (3%)	50	56	9, 23, 42, 73	0

All (87) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	533	PRO	8.9
1	C	650	SER	5.8
1	B	510	PRO	4.6
1	B	672	PRO	4.5
1	A	650	SER	4.5
1	C	673	ALA	4.4
1	C	651	GLU	4.3
1	C	674	ASP	4.2
1	A	673	ALA	4.1
1	D	683	GLY	4.1
1	D	673	ALA	4.0
1	B	674	ASP	4.0
1	A	511	ASP	3.9
1	C	562	ASN	3.8
1	D	684	SER	3.8
1	D	672	PRO	3.8
1	A	674	ASP	3.8
1	C	24	GLY	3.7
1	D	674	ASP	3.6
1	B	683	GLY	3.6
1	D	446	SER	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	371	PRO	3.5
1	C	511	ASP	3.5
1	B	684	SER	3.5
1	D	562	ASN	3.5
1	B	673	ALA	3.5
1	C	672	PRO	3.4
1	D	464	ALA	3.3
1	D	460	SER	3.3
1	D	61	ALA	3.2
1	D	531	ALA	3.1
1	C	535	ASN	3.1
1	A	649	ALA	3.1
1	C	534	GLY	3.1
1	C	531	ALA	3.0
1	C	684	SER	3.0
1	B	24	GLY	3.0
1	C	685	GLY	3.0
1	A	446	SER	2.9
1	D	557	LYS	2.9
1	D	434	GLN	2.8
1	B	446	SER	2.8
1	C	526	GLU	2.8
1	A	59	GLY	2.8
1	D	650	SER	2.8
1	D	530	SER	2.8
1	D	60	ALA	2.7
1	C	460	SER	2.7
1	C	59	GLY	2.7
1	B	371	PRO	2.7
1	A	60	ALA	2.7
1	B	511	ASP	2.7
1	A	684	SER	2.6
1	A	535	ASN	2.6
1	D	564	THR	2.6
1	D	24	GLY	2.6
1	D	212	GLY	2.6
1	D	519	ARG	2.5
1	C	446	SER	2.5
1	B	460	SER	2.5
1	A	24	GLY	2.5
1	C	60	ALA	2.5
1	C	362	GLY	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	683	GLY	2.4
1	A	672	PRO	2.4
1	A	61	ALA	2.4
1	D	369	GLY	2.4
1	D	370	GLY	2.4
1	D	384	LEU	2.3
1	D	561	HIS	2.3
1	B	562	ASN	2.3
1	A	562	ASN	2.3
1	D	535	ASN	2.3
1	B	61	ALA	2.2
1	D	435	THR	2.2
1	B	213	LYS	2.2
1	D	640	ARG	2.2
1	D	526	GLU	2.1
1	D	582	GLU	2.1
1	A	533	PRO	2.1
1	C	371	PRO	2.1
1	D	433	LYS	2.1
1	C	686	LYS	2.1
1	D	527	SER	2.1
1	B	513	ASP	2.0
1	A	600	LYS	2.0
1	D	445	VAL	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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^{th} 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^2)	Q<0.9
1	TOX	D	107	16/17	0.83	0.16	14,17,26,31	0
1	TOX	B	107	16/17	0.86	0.14	8,14,21,25	0
1	TOX	A	107	16/17	0.93	0.11	9,11,19,23	0
1	TOX	C	107	16/17	0.94	0.10	13,15,21,25	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

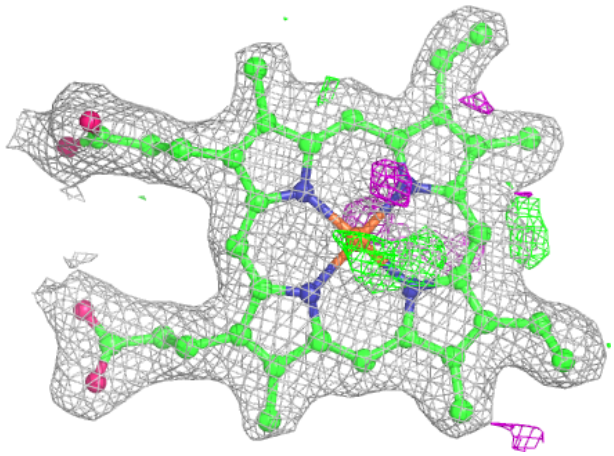
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

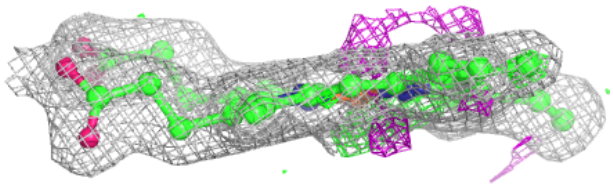
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	EPE	C	804	15/15	0.77	0.24	33,42,49,59	0
3	EPE	B	801	15/15	0.81	0.31	36,46,57,61	0
5	PGE	D	803	10/10	0.84	0.16	32,38,44,50	0
5	PGE	C	805	10/10	0.87	0.20	33,37,46,53	0
5	PGE	A	804	10/10	0.88	0.15	23,30,41,48	0
5	PGE	B	804	10/10	0.89	0.15	26,29,39,40	0
2	HEM	D	801	43/43	0.94	0.12	15,19,22,24	0
2	HEM	C	801	43/43	0.95	0.11	13,18,22,24	0
3	EPE	A	802	15/15	0.95	0.12	25,29,31,32	0
3	EPE	C	802	15/15	0.96	0.16	32,36,43,44	0
2	HEM	A	801	43/43	0.96	0.10	9,13,16,17	0
2	HEM	B	802	43/43	0.97	0.09	10,14,17,19	0
4	NA	B	803	1/1	0.97	0.10	19,19,19,19	0
4	NA	C	803	1/1	0.97	0.06	23,23,23,23	0
4	NA	D	802	1/1	0.97	0.11	26,26,26,26	0
4	NA	A	803	1/1	0.99	0.07	19,19,19,19	0

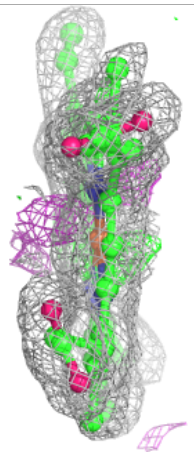
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around HEM D 801:

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

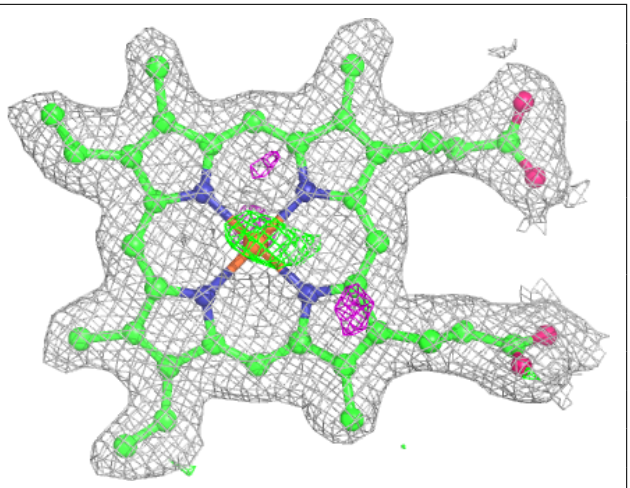


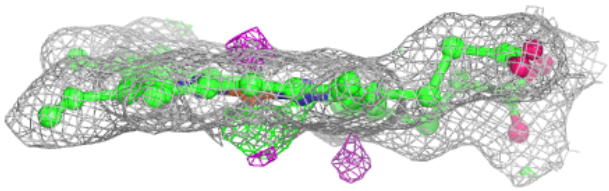


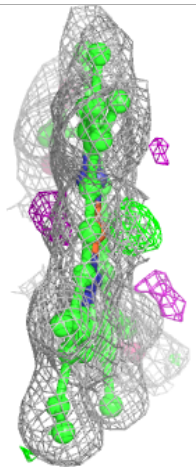


Electron density around HEM C 801:

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

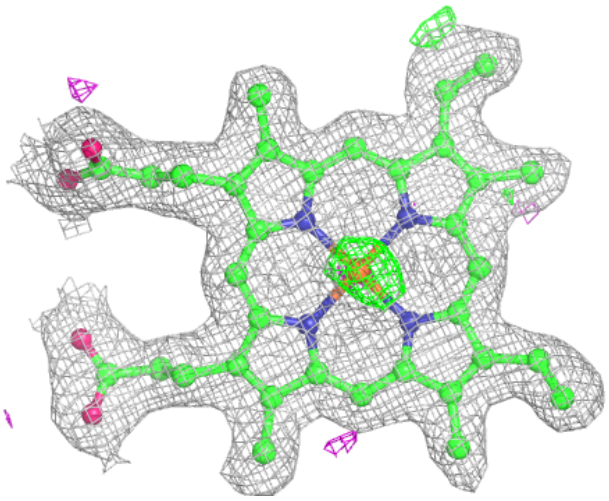


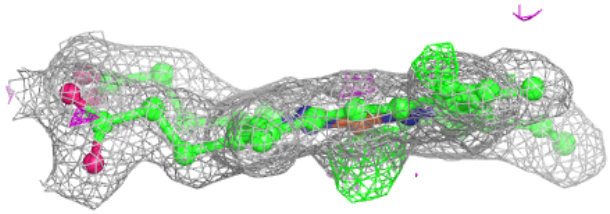


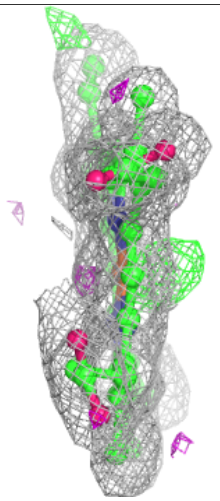


Electron density around HEM A 801:

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

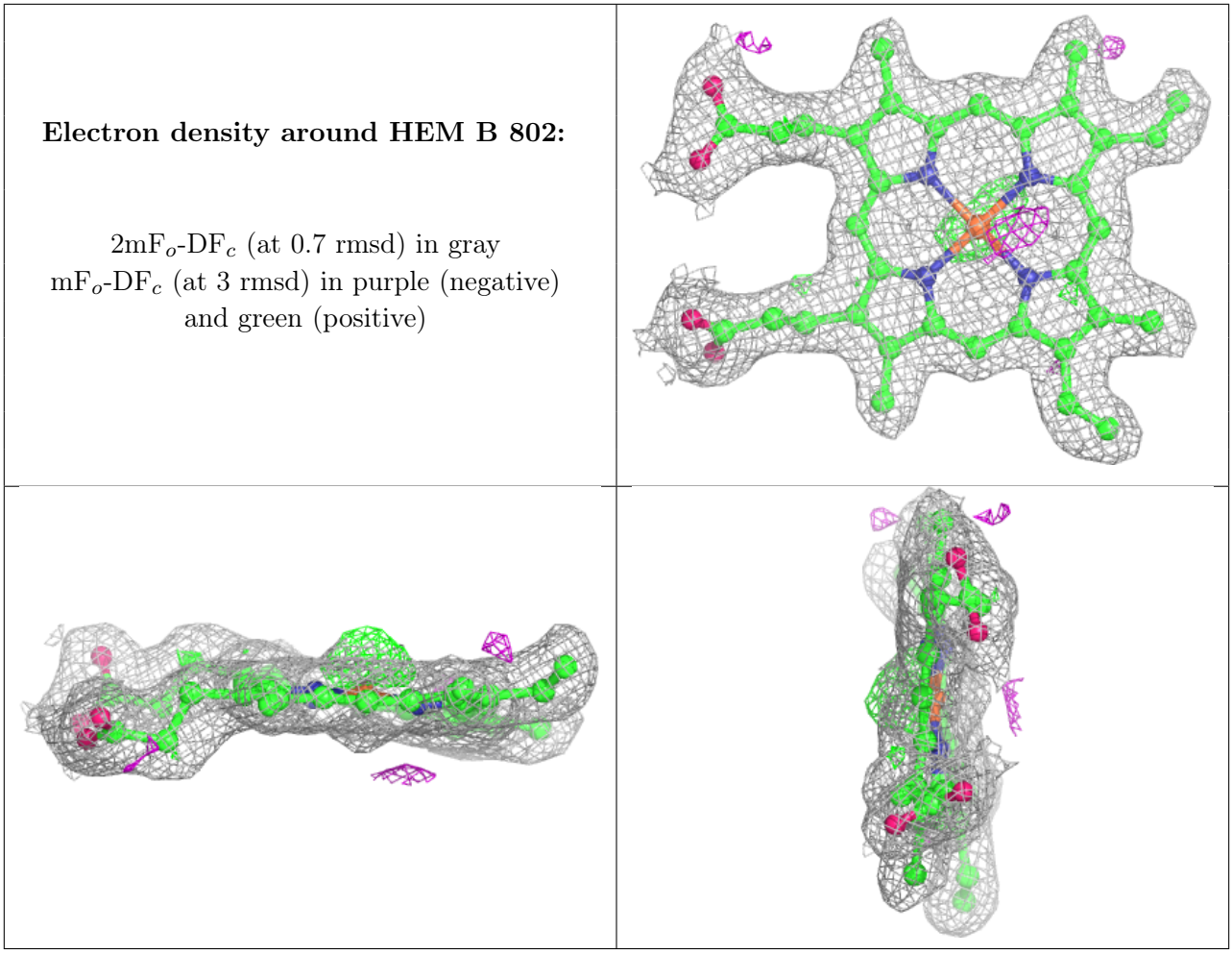






Electron density around HEM B 802:

$2mF_o$-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

There are no such residues in this entry.