Full wwPDB X-ray Structure Validation Report ⓘ

Nov 5, 2023 – 10:31 PM EST

PDB ID	:	5KQN
Title	:	Crystal structure of the H381S variant of catalase-peroxidase from B. pseudomallei
Authors	:	Loewen, P.C.
Deposited on	:	2016-07-06
Resolution	:	1.75 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:

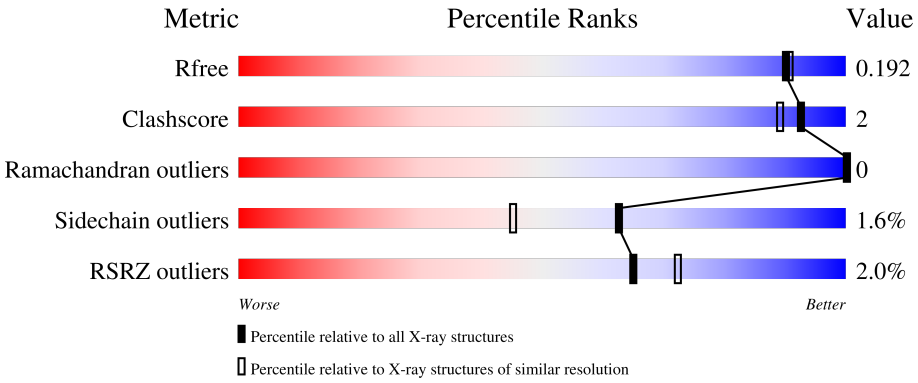
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.75 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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	728	2% 87% 9%
1	B	728	2% 88% 9%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12677 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.

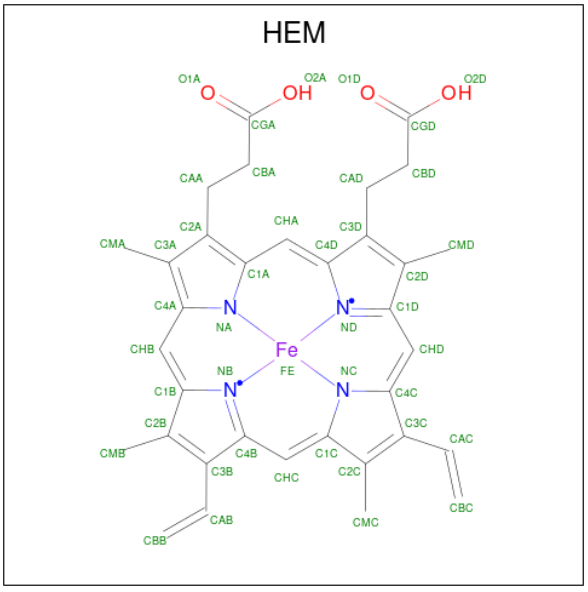
- Molecule 1 is a protein called Catalase-peroxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	713	Total	C	N	O	S	0	7	0
			5560	3508	990	1048	14			
1	B	713	Total	C	N	O	S	0	6	0
			5546	3501	987	1044	14			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	381	SER	HIS	engineered mutation	UNP Q3JNW6
B	381	SER	HIS	engineered mutation	UNP Q3JNW6

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



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

Continued on next page...

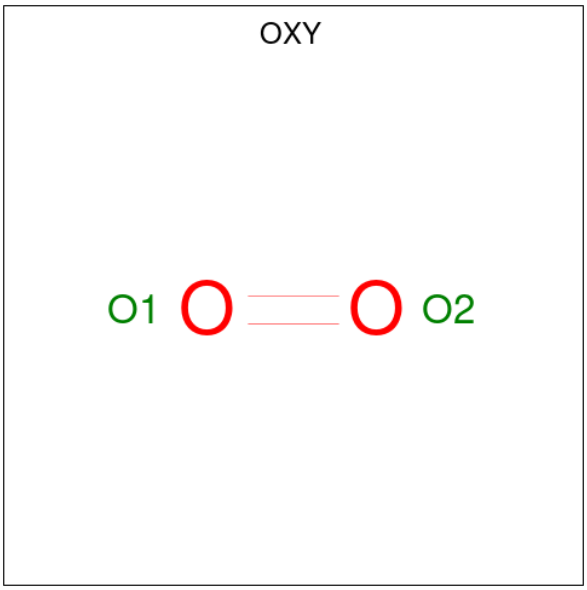
Continued from previous page...

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

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

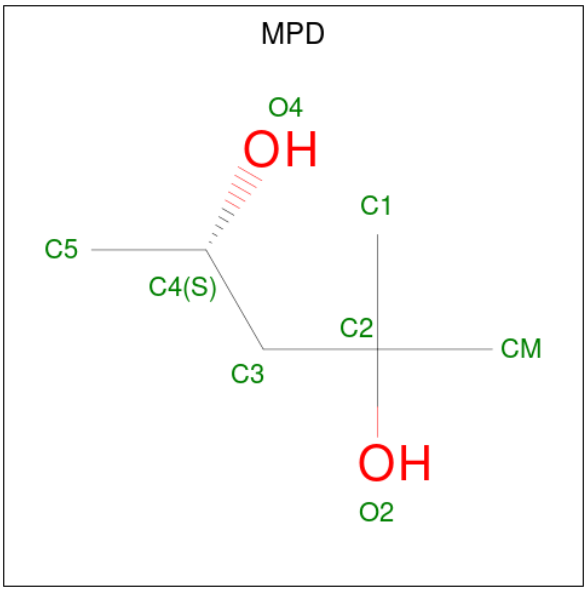
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Na	0	0
			1	1		
3	B	1	Total	Na	0	0
			1	1		

- Molecule 4 is OXYGEN MOLECULE (three-letter code: OXY) (formula: O_2).



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

- Molecule 5 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 8 6 2	0	0
5	A	1	Total C O 8 6 2	0	0
5	A	1	Total C O 8 6 2	0	0
5	B	1	Total C O 8 6 2	0	0

- Molecule 6 is water.

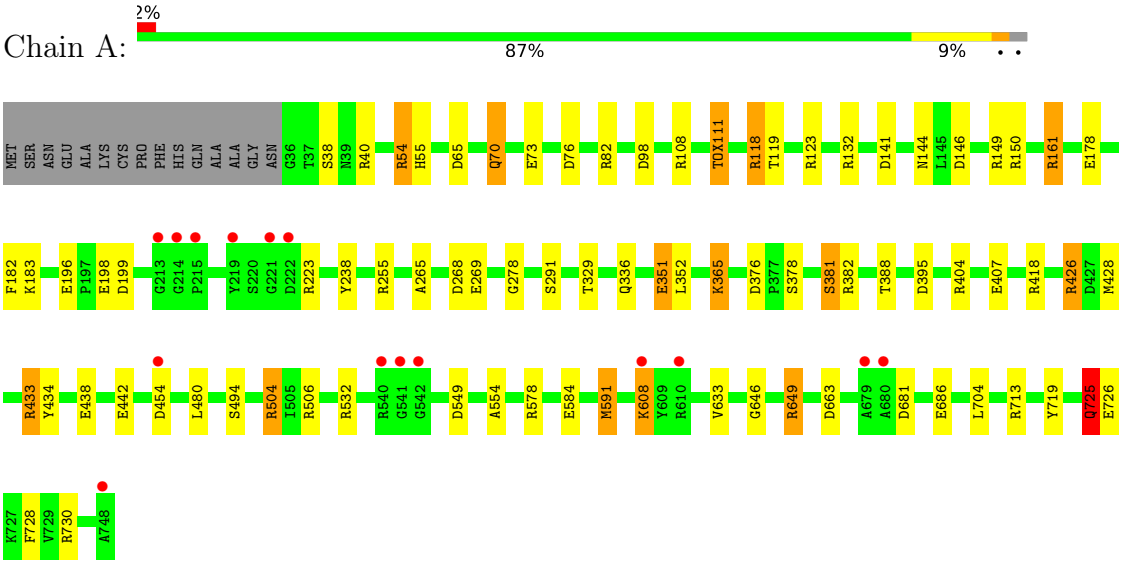
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	723	Total O 723 723	0	0
6	B	724	Total O 724 724	0	0

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.

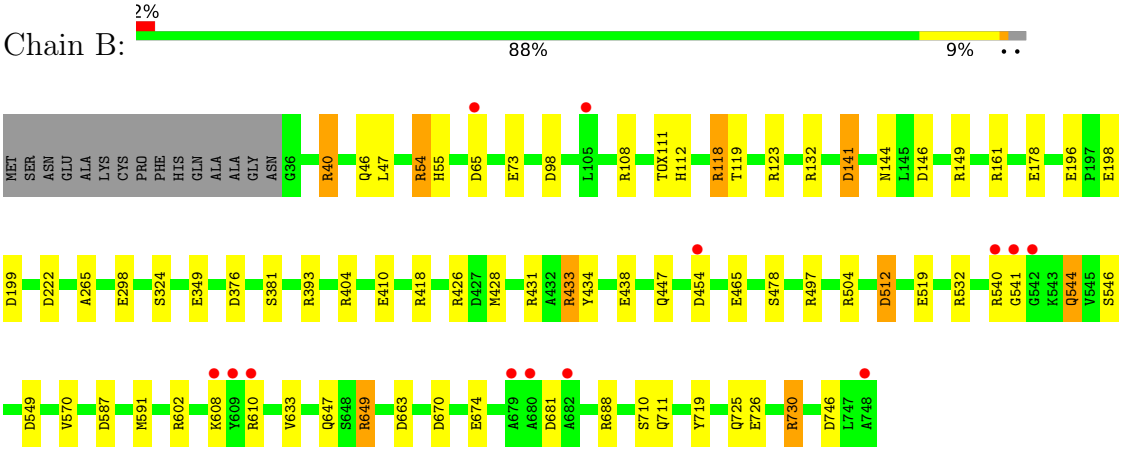
• Molecule 1: Catalase-peroxidase

Chain A:



• Molecule 1: Catalase-peroxidase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	100.99Å 114.17Å 175.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.18 – 1.75 46.18 – 1.75	Depositor EDS
% Data completeness (in resolution range)	93.9 (46.18-1.75) 93.9 (46.18-1.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	2.39 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.8.0151	Depositor
R, R_{free}	0.153 , 0.182 0.164 , 0.192	Depositor DCC
R_{free} test set	9473 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	23.5	Xtriage
Anisotropy	0.499	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 44.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	12677	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.90% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, OXY, HEM, MPD, TOX

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	1.50	37/5688 (0.7%)	1.31	56/7729 (0.7%)
1	B	1.45	29/5677 (0.5%)	1.31	49/7715 (0.6%)
All	All	1.47	66/11365 (0.6%)	1.31	105/15444 (0.7%)

All (66) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	198	GLU	CG-CD	11.07	1.68	1.51
1	A	407	GLU	CD-OE1	10.79	1.37	1.25
1	B	198	GLU	CG-CD	9.68	1.66	1.51
1	A	407	GLU	CG-CD	9.53	1.66	1.51
1	B	726	GLU	CD-OE2	9.51	1.36	1.25
1	B	410	GLU	CG-CD	9.29	1.65	1.51
1	A	381	SER	CB-OG	-8.72	1.30	1.42
1	A	725	GLN	CG-CD	8.42	1.70	1.51
1	A	442	GLU	CD-OE2	-8.21	1.16	1.25
1	B	512	ASP	CG-OD2	8.19	1.44	1.25
1	A	434	TYR	CE2-CZ	-8.06	1.28	1.38
1	A	438	GLU	CD-OE1	-7.73	1.17	1.25
1	B	512	ASP	CB-CG	7.64	1.67	1.51
1	A	686	GLU	CD-OE2	-7.52	1.17	1.25
1	A	378	SER	CB-OG	7.04	1.51	1.42
1	A	454	ASP	CB-CG	6.84	1.66	1.51
1	A	291	SER	CA-CB	6.83	1.63	1.52
1	A	178	GLU	CG-CD	6.75	1.62	1.51
1	B	746	ASP	CB-CG	6.72	1.65	1.51
1	A	223	ARG	CZ-NH2	6.42	1.41	1.33
1	A	196	GLU	CB-CG	-6.39	1.40	1.52
1	A	438	GLU	CG-CD	6.25	1.61	1.51
1	A	584	GLU	CG-CD	6.23	1.61	1.51
1	B	381	SER	CB-OG	-6.11	1.34	1.42

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	196	GLU	CG-CD	6.10	1.61	1.51
1	B	426	ARG	CZ-NH2	5.97	1.40	1.33
1	B	349	GLU	CD-OE2	-5.94	1.19	1.25
1	A	418	ARG	CZ-NH1	-5.93	1.25	1.33
1	A	426[A]	ARG	CZ-NH2	5.90	1.40	1.33
1	A	426[B]	ARG	CZ-NH2	5.90	1.40	1.33
1	B	149	ARG	CZ-NH1	-5.87	1.25	1.33
1	B	465	GLU	CD-OE1	5.82	1.32	1.25
1	B	647	GLN	CG-CD	5.80	1.64	1.51
1	B	726	GLU	CG-CD	5.79	1.60	1.51
1	A	365	LYS	N-CA	5.77	1.57	1.46
1	A	388	THR	CB-CG2	-5.66	1.33	1.52
1	B	434	TYR	CB-CG	-5.55	1.43	1.51
1	B	438	GLU	CD-OE1	-5.46	1.19	1.25
1	A	351	GLU	CD-OE2	5.40	1.31	1.25
1	B	532	ARG	CD-NE	-5.37	1.37	1.46
1	B	711	GLN	CD-NE2	-5.37	1.19	1.32
1	A	278	GLY	CA-C	5.36	1.60	1.51
1	B	46	GLN	CG-CD	5.36	1.63	1.51
1	B	602	ARG	CZ-NH2	-5.35	1.26	1.33
1	A	161	ARG	CD-NE	-5.34	1.37	1.46
1	A	269	GLU	CD-OE1	-5.32	1.19	1.25
1	A	196	GLU	CD-OE2	-5.29	1.19	1.25
1	A	73	GLU	CD-OE1	5.28	1.31	1.25
1	B	478	SER	CA-CB	5.26	1.60	1.52
1	B	73	GLU	CG-CD	5.25	1.59	1.51
1	B	544	GLN	CD-NE2	5.24	1.46	1.32
1	A	161	ARG	CG-CD	5.23	1.65	1.51
1	B	674	GLU	CG-CD	5.22	1.59	1.51
1	B	546	SER	CB-OG	5.20	1.49	1.42
1	B	710	SER	CB-OG	-5.20	1.35	1.42
1	B	298	GLU	CD-OE2	5.17	1.31	1.25
1	A	329	THR	C-O	5.16	1.33	1.23
1	B	730	ARG	CZ-NH1	5.12	1.39	1.33
1	A	646	GLY	N-CA	-5.12	1.38	1.46
1	B	196	GLU	CG-CD	5.11	1.59	1.51
1	A	494	SER	CA-CB	-5.11	1.45	1.52
1	A	73	GLU	CG-CD	5.09	1.59	1.51
1	A	584	GLU	CD-OE1	5.08	1.31	1.25
1	A	418	ARG	NE-CZ	5.05	1.39	1.33
1	B	418	ARG	CZ-NH2	-5.03	1.26	1.33
1	A	38	SER	CB-OG	5.01	1.48	1.42

All (105) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	649[A]	ARG	NE-CZ-NH2	-15.99	112.31	120.30
1	B	649[B]	ARG	NE-CZ-NH2	-15.99	112.31	120.30
1	B	512	ASP	CB-CG-OD1	-12.56	107.00	118.30
1	B	512	ASP	CB-CG-OD2	11.56	128.71	118.30
1	A	504	ARG	NE-CZ-NH2	-11.54	114.53	120.30
1	A	591	MET	CG-SD-CE	-11.24	82.22	100.20
1	B	532	ARG	NE-CZ-NH1	11.08	125.84	120.30
1	B	591	MET	CG-SD-CE	-11.05	82.51	100.20
1	B	161[A]	ARG	NE-CZ-NH2	-10.86	114.87	120.30
1	B	161[B]	ARG	NE-CZ-NH2	-10.86	114.87	120.30
1	A	150	ARG	NE-CZ-NH2	-10.85	114.88	120.30
1	B	497	ARG	NE-CZ-NH2	-10.20	115.20	120.30
1	A	395	ASP	CB-CG-OD1	9.62	126.96	118.30
1	A	149	ARG	NE-CZ-NH1	9.45	125.03	120.30
1	B	532	ARG	NE-CZ-NH2	-9.36	115.62	120.30
1	B	199	ASP	CB-CG-OD1	-9.23	109.99	118.30
1	A	532	ARG	NE-CZ-NH1	9.05	124.83	120.30
1	A	161	ARG	NE-CZ-NH1	8.88	124.74	120.30
1	B	108	ARG	NE-CZ-NH2	-8.88	115.86	120.30
1	B	541	GLY	N-CA-C	-8.76	91.21	113.10
1	A	351	GLU	OE1-CD-OE2	8.68	133.71	123.30
1	A	198	GLU	CG-CD-OE1	8.44	135.19	118.30
1	A	376	ASP	CB-CG-OD2	-8.37	110.77	118.30
1	A	255[A]	ARG	NE-CZ-NH2	-8.36	116.12	120.30
1	A	255[B]	ARG	NE-CZ-NH2	-8.36	116.12	120.30
1	A	504	ARG	NE-CZ-NH1	8.32	124.46	120.30
1	A	76	ASP	CB-CG-OD2	-8.09	111.02	118.30
1	A	532	ARG	NE-CZ-NH2	-8.08	116.26	120.30
1	A	132	ARG	NE-CZ-NH2	-8.06	116.27	120.30
1	B	198	GLU	CG-CD-OE1	8.06	134.41	118.30
1	B	376	ASP	CB-CG-OD2	-7.84	111.25	118.30
1	B	132	ARG	NE-CZ-NH1	7.63	124.12	120.30
1	A	681	ASP	CB-CG-OD2	-7.50	111.55	118.30
1	B	393	ARG	NE-CZ-NH2	-7.50	116.55	120.30
1	A	199	ASP	CB-CG-OD1	-7.36	111.68	118.30
1	A	407	GLU	OE1-CD-OE2	-7.32	114.51	123.30
1	B	123	ARG	NE-CZ-NH1	7.27	123.93	120.30
1	A	433	ARG	NE-CZ-NH1	7.26	123.93	120.30
1	A	663	ASP	CB-CG-OD2	-7.22	111.80	118.30
1	A	161	ARG	NE-CZ-NH2	-7.22	116.69	120.30
1	B	433	ARG	NE-CZ-NH1	7.18	123.89	120.30
1	A	118	ARG	NE-CZ-NH2	-7.05	116.77	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	649	ARG	NE-CZ-NH2	-7.03	116.79	120.30
1	B	40	ARG	NE-CZ-NH1	7.00	123.80	120.30
1	B	504	ARG	NE-CZ-NH2	-6.93	116.83	120.30
1	B	730	ARG	NE-CZ-NH1	6.92	123.76	120.30
1	B	663	ASP	CB-CG-OD1	6.85	124.47	118.30
1	A	730	ARG	NE-CZ-NH1	6.83	123.71	120.30
1	B	198	GLU	OE1-CD-OE2	-6.82	115.12	123.30
1	A	198	GLU	OE1-CD-OE2	-6.80	115.14	123.30
1	A	591	MET	CA-CB-CG	6.73	124.73	113.30
1	A	123	ARG	NE-CZ-NH1	6.72	123.66	120.30
1	B	118	ARG	NE-CZ-NH2	-6.69	116.96	120.30
1	A	454	ASP	CB-CG-OD2	6.64	124.28	118.30
1	A	82	ARG	NE-CZ-NH2	-6.57	117.01	120.30
1	A	681	ASP	CB-CG-OD1	6.49	124.14	118.30
1	B	108	ARG	NE-CZ-NH1	6.48	123.54	120.30
1	A	730	ARG	NE-CZ-NH2	-6.41	117.09	120.30
1	B	434	TYR	CB-CG-CD1	6.40	124.84	121.00
1	A	381	SER	N-CA-CB	6.38	120.07	110.50
1	A	132	ARG	NE-CZ-NH1	6.36	123.48	120.30
1	B	65	ASP	CB-CG-OD2	-6.13	112.78	118.30
1	A	65	ASP	CB-CG-OD2	-6.13	112.78	118.30
1	A	404	ARG	NE-CZ-NH2	-6.11	117.25	120.30
1	B	504	ARG	NE-CZ-NH1	6.07	123.33	120.30
1	A	40	ARG	NE-CZ-NH1	6.02	123.31	120.30
1	A	381	SER	CB-CA-C	-5.98	98.73	110.10
1	B	349	GLU	OE1-CD-OE2	-5.97	116.14	123.30
1	B	663	ASP	CB-CG-OD2	-5.95	112.95	118.30
1	B	222	ASP	CB-CG-OD1	-5.89	112.99	118.30
1	B	688	ARG	NE-CZ-NH2	-5.87	117.36	120.30
1	A	549	ASP	CB-CG-OD2	-5.82	113.06	118.30
1	B	404	ARG	NE-CZ-NH2	-5.77	117.41	120.30
1	B	426	ARG	NE-CZ-NH1	-5.71	117.44	120.30
1	A	404	ARG	NE-CZ-NH1	5.70	123.15	120.30
1	B	587	ASP	CB-CG-OD1	5.70	123.42	118.30
1	A	365	LYS	CA-CB-CG	5.69	125.93	113.40
1	A	70	GLN	CA-CB-CG	5.69	125.91	113.40
1	A	108	ARG	NE-CZ-NH1	5.66	123.13	120.30
1	A	578	ARG	NE-CZ-NH1	5.66	123.13	120.30
1	B	393	ARG	NE-CZ-NH1	5.64	123.12	120.30
1	B	570	VAL	CG1-CB-CG2	-5.56	102.00	110.90
1	A	407	GLU	CG-CD-OE1	5.53	129.35	118.30
1	B	549	ASP	CB-CG-OD2	-5.44	113.40	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	670	ASP	CB-CG-OD1	5.43	123.19	118.30
1	A	434	TYR	CB-CG-CD1	5.43	124.26	121.00
1	A	268	ASP	CB-CG-OD2	-5.41	113.43	118.30
1	A	704	LEU	CB-CG-CD1	5.36	120.11	111.00
1	B	681	ASP	CB-CG-OD1	5.33	123.10	118.30
1	A	728	PHE	CB-CG-CD1	-5.32	117.08	120.80
1	A	713	ARG	NE-CZ-NH2	-5.29	117.66	120.30
1	A	183	LYS	CD-CE-NZ	-5.25	99.62	111.70
1	B	454	ASP	CB-CG-OD1	5.25	123.02	118.30
1	B	98	ASP	CB-CG-OD2	5.21	122.99	118.30
1	B	65	ASP	CB-CG-OD1	5.21	122.98	118.30
1	B	404	ARG	NE-CZ-NH1	5.17	122.88	120.30
1	B	649[A]	ARG	NH1-CZ-NH2	5.14	125.06	119.40
1	B	649[B]	ARG	NH1-CZ-NH2	5.14	125.06	119.40
1	A	506	ARG	NE-CZ-NH2	-5.09	117.75	120.30
1	B	540	ARG	NE-CZ-NH1	5.09	122.84	120.30
1	A	182	PHE	CB-CG-CD1	5.08	124.36	120.80
1	A	150	ARG	NE-CZ-NH1	5.08	122.84	120.30
1	B	730	ARG	NE-CZ-NH2	-5.08	117.76	120.30
1	A	506	ARG	NE-CZ-NH1	5.05	122.83	120.30
1	A	98	ASP	CB-CG-OD2	5.01	122.81	118.30

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	5560	0	5371	16	0
1	B	5546	0	5366	15	0
2	A	43	0	30	0	0
2	B	43	0	30	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	2	0	0	0	0
4	B	2	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	24	0	42	1	0
5	B	8	0	14	2	0
6	A	723	0	0	4	0
6	B	724	0	0	8	0
All	All	12677	0	10853	34	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 (34) 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:119[B]:THR:HG21	6:A:999:HOH:O	1.82	0.80
1:B:119[B]:THR:HG21	6:B:1296:HOH:O	1.84	0.77
1:B:512:ASP:OD1	6:B:901:HOH:O	2.08	0.70
1:B:519:GLU:OE1	6:B:902:HOH:O	2.09	0.70
1:A:608:LYS:O	1:A:608:LYS:HD2	2.05	0.56
1:B:178:GLU:OE1	6:B:904:HOH:O	2.18	0.55
1:A:336[B]:GLN:HG3	6:A:1057:HOH:O	2.07	0.54
1:A:119[A]:THR:CG2	1:A:265:ALA:HB2	2.38	0.54
1:A:504:ARG:HD2	6:A:1082:HOH:O	2.10	0.52
1:A:365:LYS:HA	1:A:382:ARG:HH11	1.75	0.52
1:A:633[A]:VAL:CG2	1:A:719:TYR:CZ	2.94	0.51
5:A:805:MPD:H52	5:A:805:MPD:H11	1.92	0.51
5:B:804:MPD:O4	5:B:804:MPD:H12	2.12	0.50
4:B:803:OXY:O2	6:B:903:HOH:O	2.18	0.49
1:B:428:MET:O	1:B:433:ARG:HD3	2.13	0.48
1:B:54:ARG:HB3	1:B:55:HIS:CD2	2.49	0.47
1:B:633[A]:VAL:CG2	1:B:719:TYR:CZ	2.98	0.47
1:A:54:ARG:HB3	1:A:55:HIS:CD2	2.50	0.47
1:B:47:LEU:HB2	6:B:908:HOH:O	2.16	0.46
1:A:70:GLN:NE2	6:A:921:HOH:O	2.49	0.46
1:A:725:GLN:NE2	1:A:726[A]:GLU:HG3	2.30	0.46
1:A:426[B]:ARG:HA	1:A:426[B]:ARG:HD2	1.86	0.44
1:B:324:SER:O	5:B:804:MPD:H53	2.17	0.44
1:A:428:MET:O	1:A:433:ARG:HD3	2.18	0.44
1:B:725:GLN:CD	6:B:1288:HOH:O	2.57	0.43
1:B:512:ASP:HB2	6:B:1336:HOH:O	2.18	0.43
1:B:431:ARG:HD2	1:B:447:GLN:OE1	2.18	0.43
1:B:119[A]:THR:CG2	1:B:265:ALA:HB2	2.49	0.42
1:B:144:ASN:HA	1:B:146:ASP:OD1	2.20	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:480:LEU:HD22	1:A:554:ALA:HB1	2.01	0.41
1:B:112:HIS:CE1	1:B:141:ASP:O	2.73	0.41
1:A:351:GLU:HG3	1:A:352:LEU:N	2.36	0.41
1:A:144:ASN:HA	1:A:146:ASP:OD1	2.21	0.40
1:A:111:TOX:H9	1:A:238:TYR:OH	2.21	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	717/728 (98%)	706 (98%)	11 (2%)	0	100	100
1	B	716/728 (98%)	703 (98%)	13 (2%)	0	100	100
All	All	1433/1456 (98%)	1409 (98%)	24 (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	556/560 (99%)	547 (98%)	9 (2%)	62	45
1	B	555/560 (99%)	545 (98%)	10 (2%)	59	40
All	All	1111/1120 (99%)	1092 (98%)	19 (2%)	62	42

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	54	ARG
1	A	118	ARG
1	A	141	ASP
1	A	161	ARG
1	A	381	SER
1	A	591	MET
1	A	608	LYS
1	A	649	ARG
1	A	725	GLN
1	B	40	ARG
1	B	54	ARG
1	B	118	ARG
1	B	141	ASP
1	B	544	GLN
1	B	608	LYS
1	B	610	ARG
1	B	649[A]	ARG
1	B	649[B]	ARG
1	B	730	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	247	ASN
1	B	227	ASN
1	B	406	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)

2 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	111	1	10,17,18	1.87	3 (30%)	10,23,25	3.19	5 (50%)
1	TOX	B	111	2,1	10,17,18	2.72	5 (50%)	10,23,25	1.82	4 (40%)

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	111	1	-	2/4/8/10	0/2/2/2
1	TOX	B	111	2,1	-	2/4/8/10	0/2/2/2

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	111	TOX	CD1-NE1	-4.91	1.34	1.39
1	A	111	TOX	CE3-CD2	-4.07	1.33	1.42
1	B	111	TOX	O-C	3.66	1.34	1.19
1	B	111	TOX	CE3-CD2	-3.19	1.35	1.42
1	B	111	TOX	CB-CA	3.07	1.60	1.53
1	B	111	TOX	CZ2-CE2	-2.71	1.35	1.41
1	A	111	TOX	CH2-CZ3	2.24	1.43	1.38
1	A	111	TOX	O-C	2.01	1.27	1.19

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	111	TOX	CB-CG-CD1	-6.09	120.44	127.97
1	A	111	TOX	CB-CG-CD2	5.91	135.45	126.25
1	B	111	TOX	CB-CG-CD1	-3.47	123.68	127.97
1	A	111	TOX	CB-CA-C	3.21	117.48	111.47
1	A	111	TOX	CH2-CZ3-CE3	-2.66	116.71	120.44
1	B	111	TOX	CB-CG-CD2	2.54	130.19	126.25
1	B	111	TOX	CH2-CZ3-CE3	-2.52	116.90	120.44
1	A	111	TOX	CZ3-CH2-CZ2	2.24	123.58	120.44
1	B	111	TOX	CB-CA-C	2.15	115.50	111.47

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	111	TOX	N-CA-CB-CG
1	B	111	TOX	N-CA-CB-CG
1	B	111	TOX	C-CA-CB-CG
1	A	111	TOX	C-CA-CB-CG

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	111	TOX	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 2 are monoatomic - leaving 8 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	$\#\|Z\| > 2$
4	OXY	B	803	-	1,1,1	0.91	0	-		
2	HEM	B	801	1	41,50,50	1.83	8 (19%)	45,82,82	1.51	10 (22%)
4	OXY	A	803	-	1,1,1	0.50	0	-		
5	MPD	A	805	-	7,7,7	0.64	0	9,10,10	2.11	3 (33%)
5	MPD	A	806	-	7,7,7	0.73	0	9,10,10	1.21	1 (11%)
5	MPD	B	804	-	7,7,7	0.53	0	9,10,10	1.63	3 (33%)
5	MPD	A	804	-	7,7,7	0.65	0	9,10,10	1.54	2 (22%)
2	HEM	A	801	1	41,50,50	1.14	2 (4%)	45,82,82	2.07	17 (37%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEM	B	801	1	-	3/12/54/54	-
5	MPD	A	806	-	-	1/5/5/5	-
5	MPD	B	804	-	-	3/5/5/5	-
5	MPD	A	805	-	-	3/5/5/5	-
5	MPD	A	804	-	-	0/5/5/5	-
2	HEM	A	801	1	-	3/12/54/54	-

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	801	HEM	C3C-C2C	-5.92	1.32	1.40
2	B	801	HEM	CMD-C2D	-3.88	1.42	1.50
2	B	801	HEM	C4D-ND	-3.54	1.34	1.40
2	A	801	HEM	C4D-ND	-3.40	1.34	1.40
2	B	801	HEM	C1B-NB	-3.07	1.35	1.40
2	B	801	HEM	C1B-C2B	-2.94	1.38	1.44
2	B	801	HEM	O2D-CGD	-2.49	1.22	1.30
2	B	801	HEM	C1A-NA	2.23	1.40	1.36
2	A	801	HEM	CAD-C3D	2.03	1.56	1.51
2	B	801	HEM	C4D-C3D	2.01	1.48	1.45

All (36) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	801	HEM	C4A-C3A-C2A	-6.85	102.23	107.00
2	A	801	HEM	C3B-C2B-C1B	-4.81	102.92	106.49
5	A	805	MPD	O2-C2-CM	-4.31	94.24	108.08
2	B	801	HEM	CMA-C3A-C4A	-3.72	122.75	128.46
2	B	801	HEM	CHC-C4B-NB	3.70	128.46	124.43
2	A	801	HEM	C2B-C1B-NB	3.10	113.51	109.84
2	A	801	HEM	O2A-CGA-CBA	3.06	123.85	114.03
2	A	801	HEM	C3C-C4C-NC	-3.05	105.18	110.94
2	A	801	HEM	C4B-C3B-C2B	3.02	109.52	107.11
2	B	801	HEM	O1A-CGA-CBA	-2.95	113.59	123.08
2	A	801	HEM	CHD-C1D-ND	2.94	127.62	124.43
2	B	801	HEM	C1B-NB-C4B	2.88	108.05	105.07
5	A	805	MPD	CM-C2-C1	2.86	116.53	110.57
2	A	801	HEM	O1A-CGA-CBA	-2.84	113.97	123.08
5	A	804	MPD	O2-C2-C1	2.78	117.01	108.08

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	804	MPD	O2-C2-CM	-2.74	99.28	108.08
2	B	801	HEM	C4B-CHC-C1C	2.73	126.17	122.56
2	A	801	HEM	CAA-CBA-CGA	-2.73	106.10	113.76
5	B	804	MPD	O2-C2-C3	-2.65	99.83	109.80
2	A	801	HEM	CBA-CAA-C2A	2.64	117.13	112.62
2	A	801	HEM	CMA-C3A-C4A	2.61	132.47	128.46
5	A	806	MPD	O2-C2-CM	-2.60	99.73	108.08
2	A	801	HEM	CMC-C2C-C3C	2.55	129.45	124.68
5	B	804	MPD	C1-C2-C3	2.53	121.75	109.96
2	A	801	HEM	CBD-CAD-C3D	2.50	119.58	112.63
5	A	805	MPD	O4-C4-C3	-2.50	101.27	111.36
2	B	801	HEM	CAA-CBA-CGA	-2.48	106.80	113.76
2	A	801	HEM	C1B-NB-C4B	-2.34	102.66	105.07
5	B	804	MPD	O2-C2-CM	2.30	115.47	108.08
2	B	801	HEM	CMA-C3A-C2A	2.23	129.14	124.94
2	B	801	HEM	C2C-C3C-C4C	2.19	108.43	106.90
2	A	801	HEM	CHB-C1B-NB	-2.14	121.73	124.38
2	B	801	HEM	C4C-CHD-C1D	2.11	125.34	122.56
2	A	801	HEM	C1D-C2D-C3D	-2.09	104.76	106.96
2	B	801	HEM	CBA-CAA-C2A	2.08	116.18	112.62
2	A	801	HEM	C2C-C3C-C4C	2.07	108.35	106.90

There are no chirality outliers.

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	805	MPD	C2-C3-C4-O4
5	B	804	MPD	C1-C2-C3-C4
5	B	804	MPD	O2-C2-C3-C4
5	A	805	MPD	C2-C3-C4-C5
5	A	805	MPD	CM-C2-C3-C4
5	B	804	MPD	CM-C2-C3-C4
2	B	801	HEM	CAA-CBA-CGA-O1A
2	B	801	HEM	CAA-CBA-CGA-O2A
2	A	801	HEM	CAA-CBA-CGA-O1A
2	A	801	HEM	CAA-CBA-CGA-O2A
2	A	801	HEM	CAD-CBD-CGD-O2D
5	A	806	MPD	C2-C3-C4-O4
2	B	801	HEM	CAD-CBD-CGD-O2D

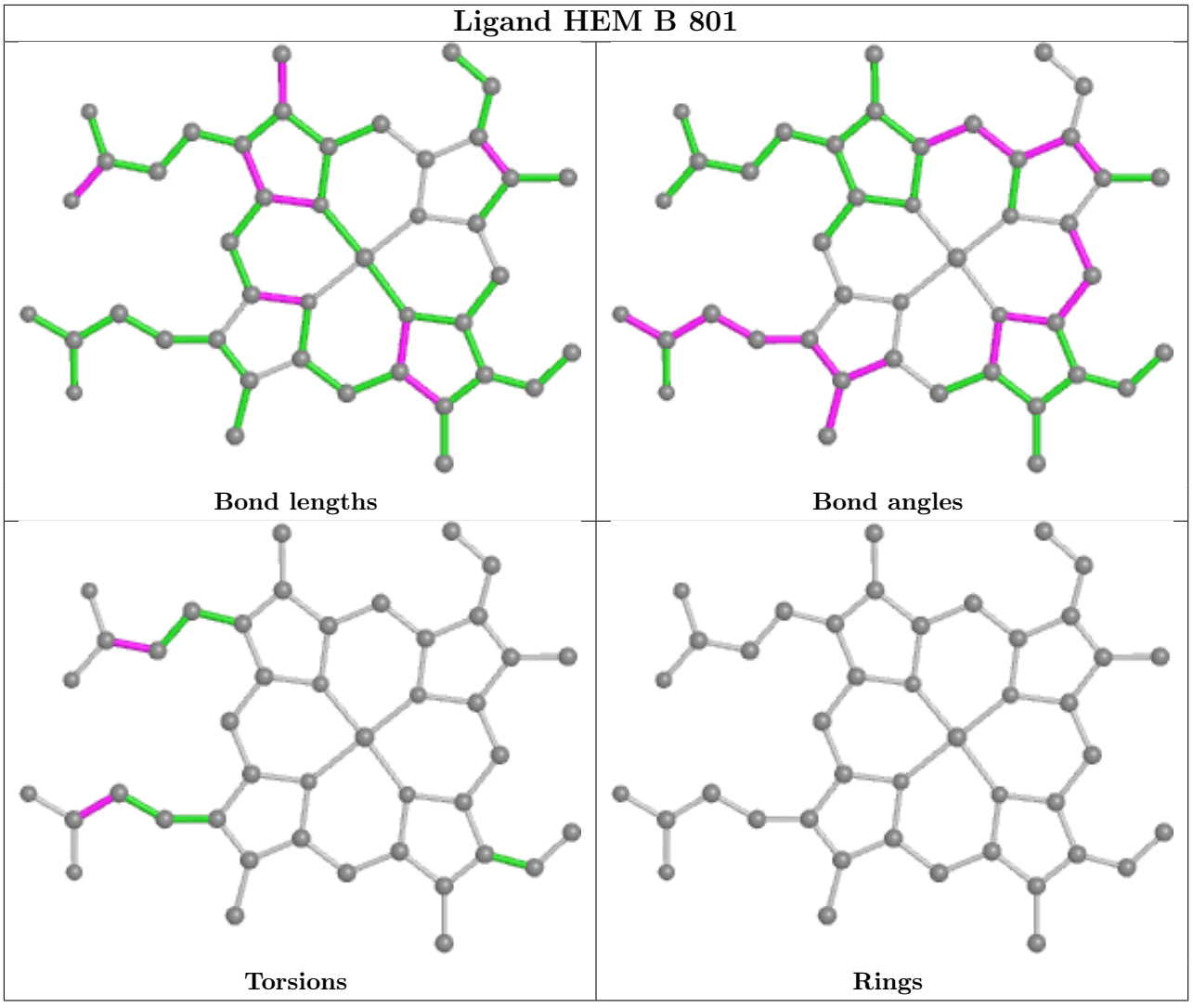
There are no ring outliers.

3 monomers are involved in 4 short contacts:

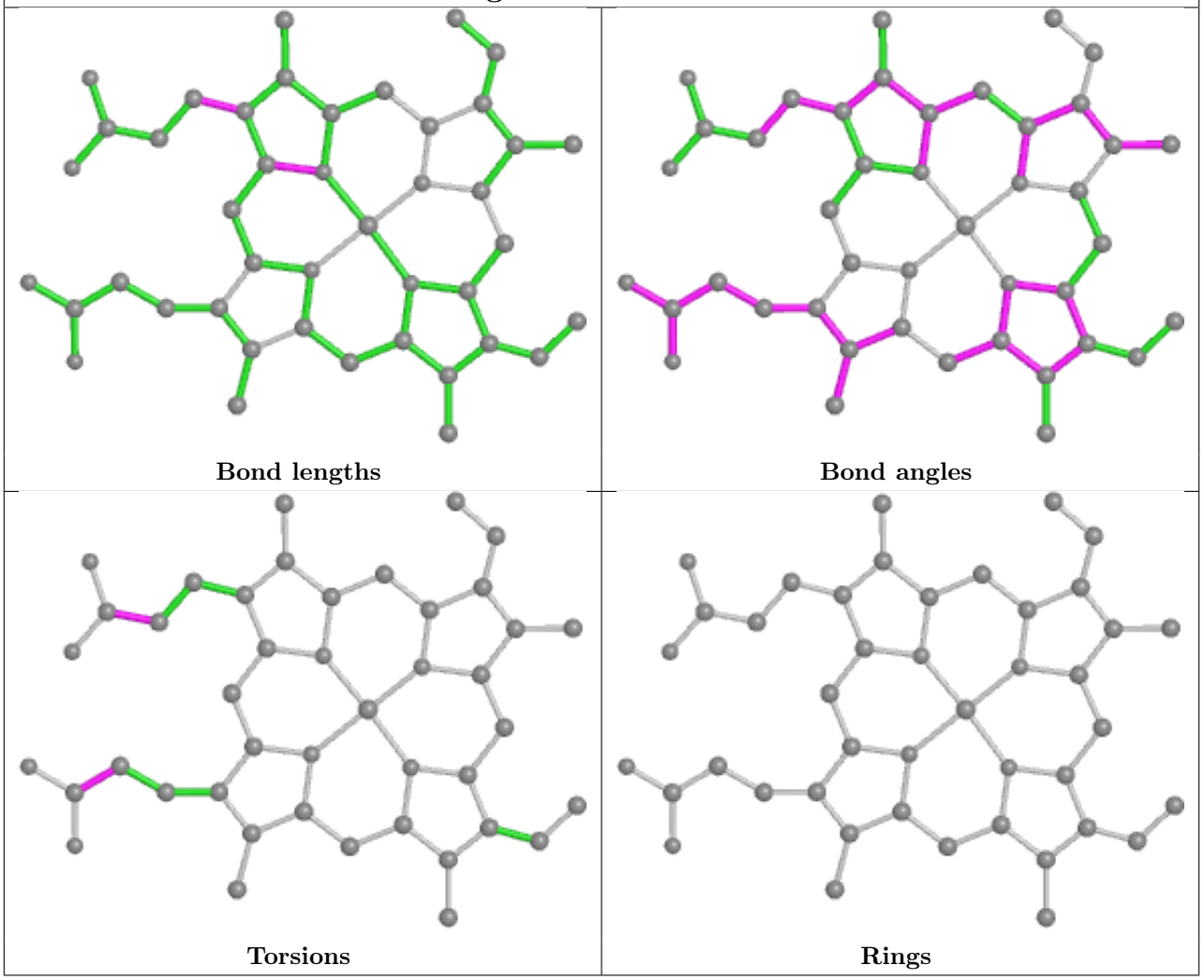
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	803	OXY	1	0
5	A	805	MPD	1	0
5	B	804	MPD	2	0

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 B 801



Ligand HEM A 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	712/728 (97%)	-0.32	15 (2%)	63	71	17, 25, 43, 81	0
1	B	712/728 (97%)	-0.40	13 (1%)	68	76	18, 24, 42, 82	0
All	All	1424/1456 (97%)	-0.36	28 (1%)	65	72	17, 25, 43, 82	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	748	ALA	5.7
1	A	541	GLY	5.3
1	B	540	ARG	5.1
1	B	748	ALA	4.5
1	A	540	ARG	3.8
1	B	610	ARG	3.8
1	A	608	LYS	3.7
1	B	680	ALA	3.0
1	B	679	ALA	2.9
1	A	221	GLY	2.9
1	A	222	ASP	2.9
1	A	454	ASP	2.9
1	A	215	PRO	2.7
1	B	542	GLY	2.7
1	A	679	ALA	2.7
1	B	541	GLY	2.6
1	A	610	ARG	2.6
1	A	213	GLY	2.6
1	A	542	GLY	2.5
1	B	454	ASP	2.4
1	A	680	ALA	2.3
1	B	682	ALA	2.2
1	B	105	LEU	2.2
1	B	609	TYR	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	214	GLY	2.1
1	B	608	LYS	2.1
1	A	219	TYR	2.1
1	B	65	ASP	2.1

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	A	111	16/17	0.98	0.10	16,19,26,30	2
1	TOX	B	111	16/17	0.98	0.11	17,19,31,33	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.

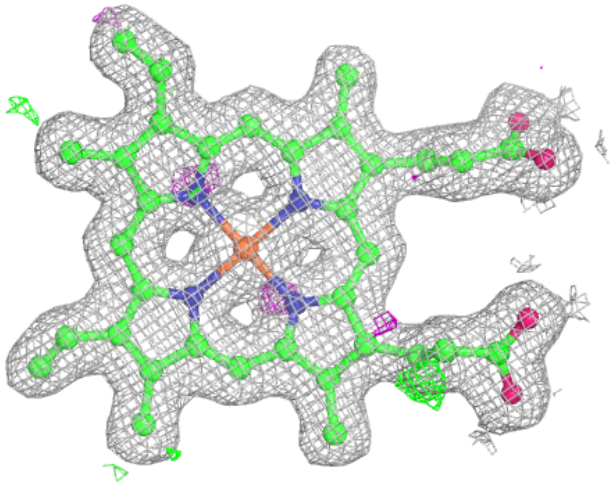
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	MPD	A	804	8/8	0.80	0.18	58,67,73,73	0
5	MPD	A	806	8/8	0.81	0.30	73,80,84,85	0
5	MPD	B	804	8/8	0.90	0.13	53,60,68,68	0
4	OXY	B	803	2/2	0.91	0.14	36,36,36,37	0
4	OXY	A	803	2/2	0.95	0.10	34,34,34,45	0
5	MPD	A	805	8/8	0.95	0.15	39,45,66,68	0
2	HEM	A	801	43/43	0.98	0.09	17,19,22,26	0
2	HEM	B	801	43/43	0.99	0.12	18,19,21,27	0
3	NA	A	802	1/1	0.99	0.04	20,20,20,20	0
3	NA	B	802	1/1	0.99	0.04	21,21,21,21	0

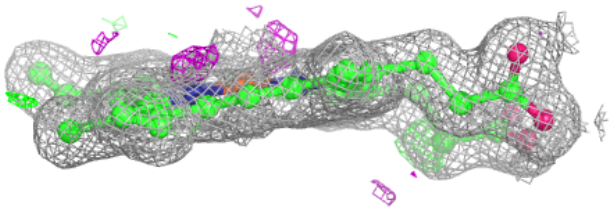
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers

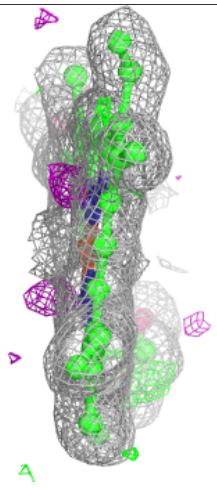
as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around HEM A 801:

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

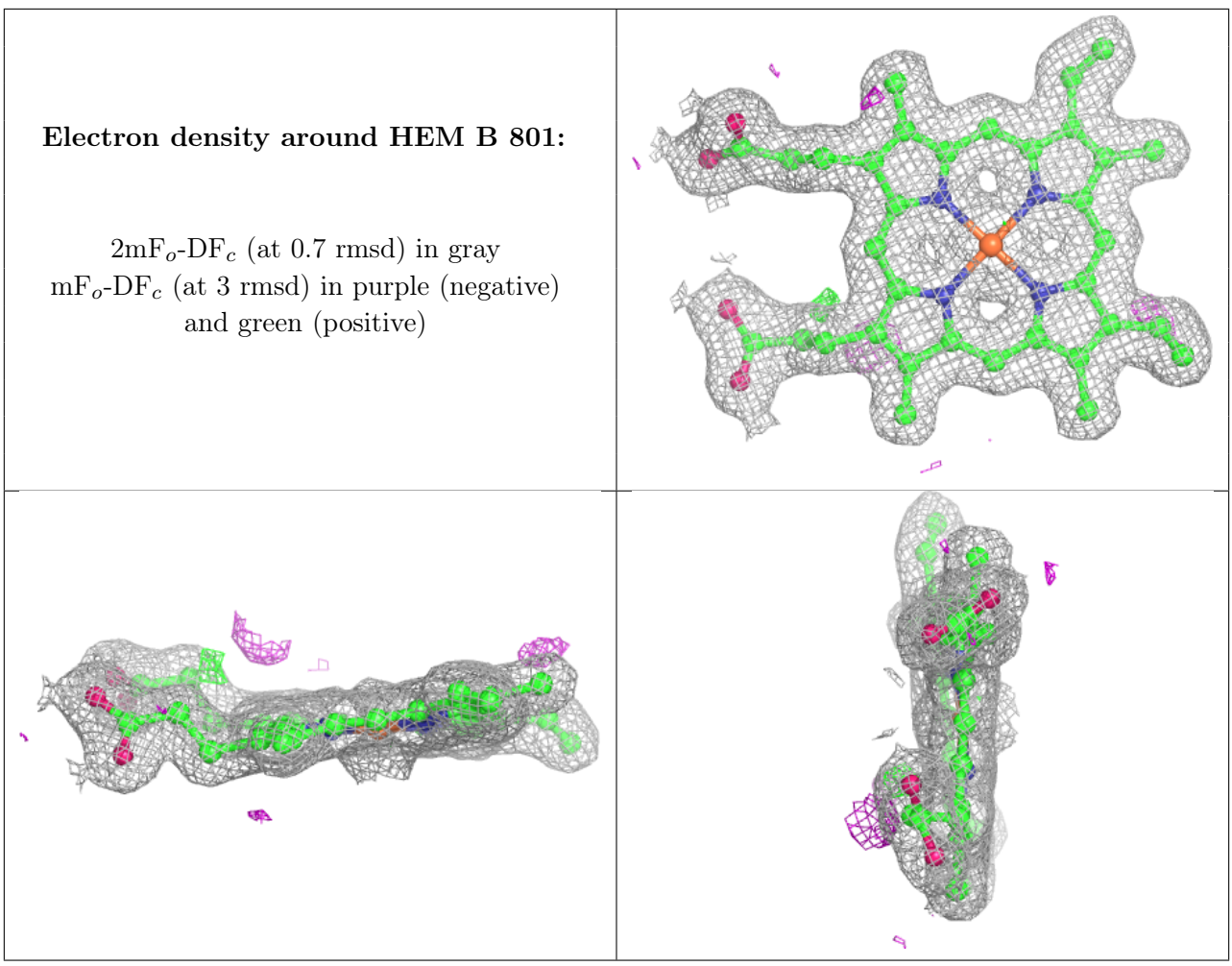






Electron density around HEM B 801:

$2mF_o$-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.