Full wwPDB X-ray Structure Validation Report (i)

Nov 7, 2023 – 05:27 AM EST

PDB ID	:	5SYW
Title	:	Crystal structure of Burkhoderia pseudomallei KatG variant Q233E
Authors	:	Loewen, P.C.
Deposited on	:	2016-08-12
Resolution	:	1.85 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

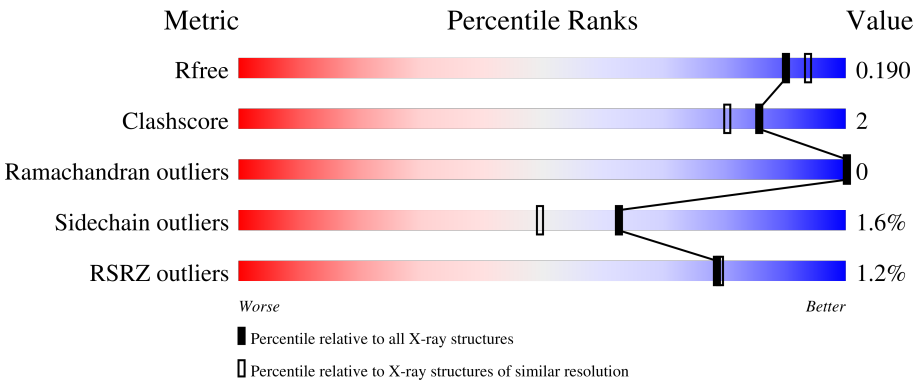
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

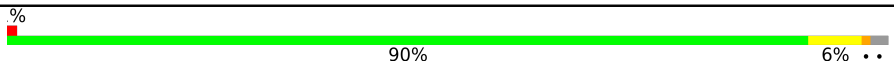
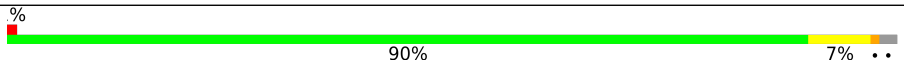
The reported resolution of this entry is 1.85 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=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	
1	B	728	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	MPD	B	806	-	-	X	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 12680 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.

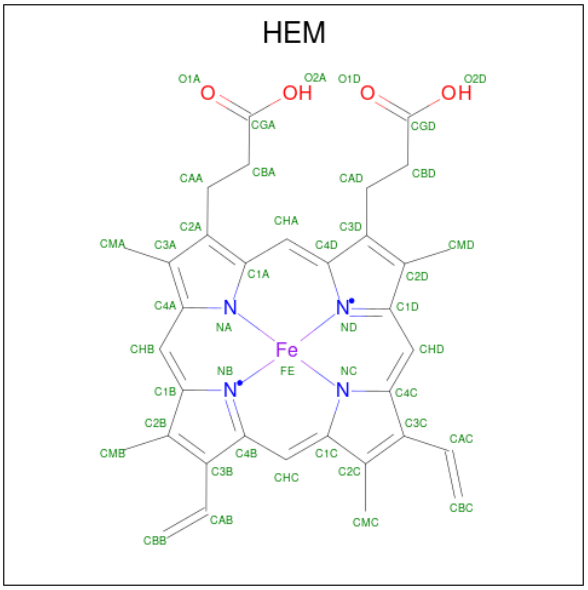
- Molecule 1 is a protein called Catalase-peroxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	713	5534	3494	986	1040	14	0	5	0
1	B	713	5544	3503	986	1041	14	0	8	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	233	GLU	GLN	engineered mutation	UNP Q3JNW6
B	233	GLU	GLN	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
			Total	C	Fe	N	O		
2	A	1	43	34	1	4	4	0	0

Continued on next page...

Continued from previous page...

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

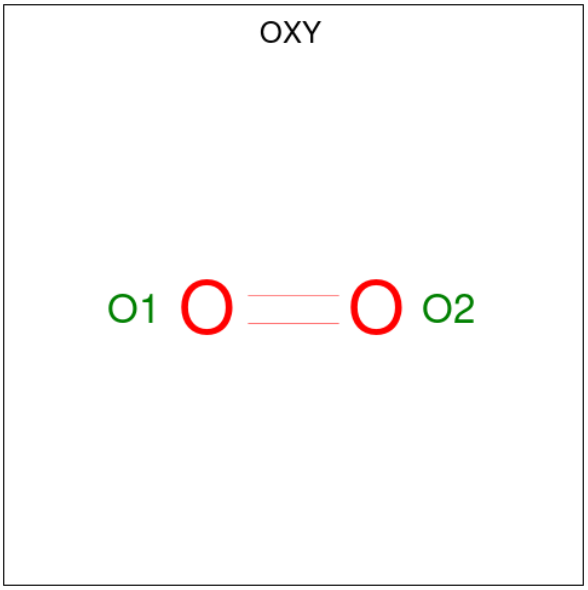
- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

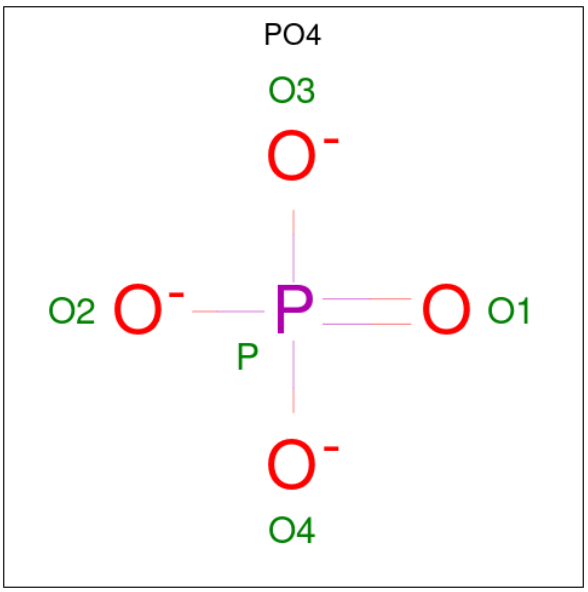
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
4	A	1	1	1	0	0
4	B	1	1	1	0	0

- Molecule 5 is OXYGEN MOLECULE (three-letter code: OXY) (formula: O_2).



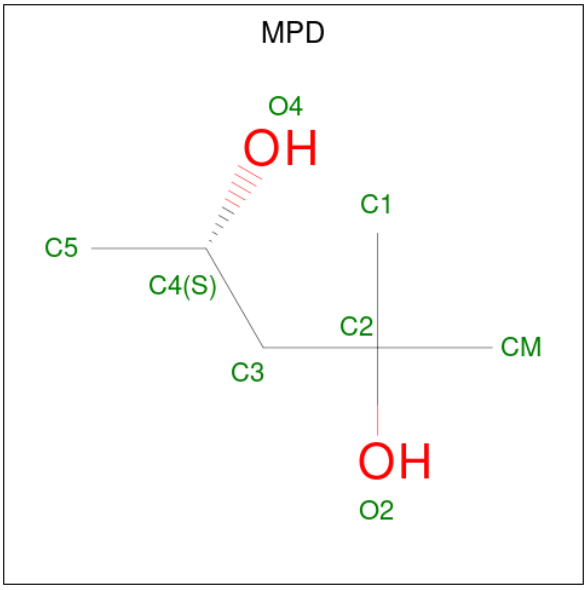
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	1	2	2	0	0
5	B	1	2	2	0	0

- Molecule 6 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O P 5 4 1	0	0
6	B	1	Total O P 5 4 1	0	0

- Molecule 7 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



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

- Molecule 8 is water.

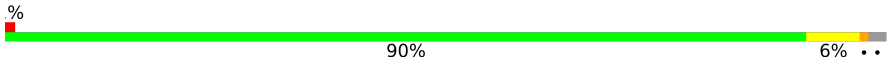
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	743	743	743	0	0
			Total	O		
8	B	715	715	715	0	0

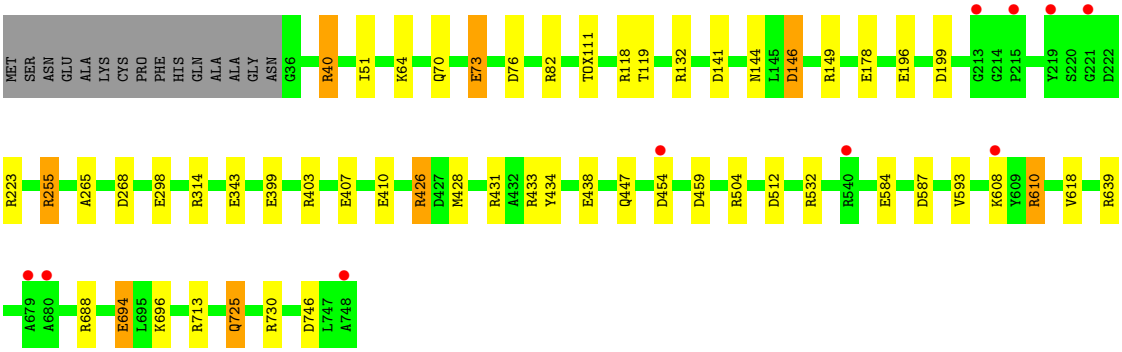
3 Residue-property plots (i)

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

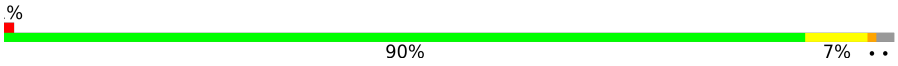
Chain A:

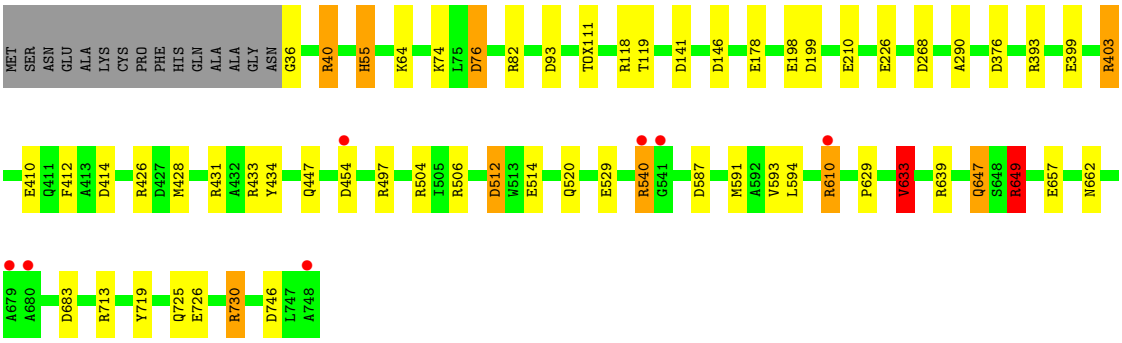




• Molecule 1: Catalase-peroxidase

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	101.02Å 114.19Å 174.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.85 47.81 – 1.85	Depositor EDS
% Data completeness (in resolution range)	98.8 (20.00-1.85) 98.8 (47.81-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.87 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.8.0151	Depositor
R, R_{free}	0.149 , 0.182 0.161 , 0.190	Depositor DCC
R_{free} test set	8419 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	21.0	Xtriage
Anisotropy	0.751	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 44.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	12680	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.86% 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: CL, PO4, HEM, TOX, MPD, OXY, 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	1.26	17/5668 (0.3%)	1.12	30/7702 (0.4%)
1	B	1.25	18/5689 (0.3%)	1.16	35/7731 (0.5%)
All	All	1.25	35/11357 (0.3%)	1.14	65/15433 (0.4%)

All (35) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	198	GLU	CD-OE2	8.80	1.35	1.25
1	B	512	ASP	CG-OD2	8.49	1.44	1.25
1	A	407	GLU	CG-CD	7.83	1.63	1.51
1	B	726	GLU	CD-OE1	7.75	1.34	1.25
1	B	512	ASP	CB-CG	7.50	1.67	1.51
1	B	746	ASP	CB-CG	6.81	1.66	1.51
1	B	434	TYR	CG-CD1	-6.75	1.30	1.39
1	A	178	GLU	CG-CD	6.65	1.61	1.51
1	A	407	GLU	CD-OE1	6.65	1.32	1.25
1	A	410	GLU	CD-OE2	6.62	1.32	1.25
1	B	210	GLU	CD-OE2	6.62	1.32	1.25
1	A	426[A]	ARG	CZ-NH2	6.59	1.41	1.33
1	A	426[B]	ARG	CZ-NH2	6.59	1.41	1.33
1	A	725	GLN	CG-CD	6.46	1.66	1.51
1	A	407	GLU	CD-OE2	6.24	1.32	1.25
1	A	399	GLU	CD-OE1	6.15	1.32	1.25
1	B	529	GLU	CG-CD	6.04	1.61	1.51
1	A	196	GLU	CG-CD	5.96	1.60	1.51
1	A	298	GLU	CG-CD	5.93	1.60	1.51
1	A	434	TYR	CG-CD1	-5.93	1.31	1.39
1	B	178	GLU	CG-CD	5.93	1.60	1.51
1	B	657	GLU	CD-OE1	5.75	1.31	1.25
1	A	584	GLU	CG-CD	5.65	1.60	1.51
1	B	36	GLY	N-CA	5.46	1.54	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	226	GLU	CG-CD	5.43	1.60	1.51
1	B	657	GLU	CG-CD	5.41	1.60	1.51
1	A	532	ARG	NE-CZ	-5.41	1.26	1.33
1	A	73	GLU	CG-CD	5.39	1.60	1.51
1	B	514	GLU	CD-OE2	-5.26	1.19	1.25
1	A	223	ARG	CZ-NH1	5.20	1.39	1.33
1	B	726	GLU	CG-CD	5.16	1.59	1.51
1	B	410	GLU	CG-CD	5.09	1.59	1.51
1	A	76	ASP	CB-CG	5.05	1.62	1.51
1	B	226	GLU	CB-CG	-5.04	1.42	1.52
1	B	410	GLU	CD-OE1	5.00	1.31	1.25

All (65) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	512	ASP	CB-CG-OD1	-11.40	108.04	118.30
1	B	512	ASP	CB-CG-OD2	10.69	127.92	118.30
1	B	633[A]	VAL	CG1-CB-CG2	-8.79	96.84	110.90
1	B	633[B]	VAL	CG1-CB-CG2	-8.79	96.84	110.90
1	B	639	ARG	NE-CZ-NH1	8.19	124.39	120.30
1	A	639	ARG	NE-CZ-NH1	7.97	124.28	120.30
1	B	403	ARG	NE-CZ-NH1	7.92	124.26	120.30
1	B	730[A]	ARG	NE-CZ-NH2	-7.82	116.39	120.30
1	B	730[B]	ARG	NE-CZ-NH2	-7.82	116.39	120.30
1	A	426[A]	ARG	NE-CZ-NH1	-7.81	116.40	120.30
1	A	426[B]	ARG	NE-CZ-NH1	-7.81	116.40	120.30
1	A	76	ASP	CB-CG-OD2	-7.77	111.31	118.30
1	B	414	ASP	CB-CG-OD2	7.72	125.25	118.30
1	A	82	ARG	NE-CZ-NH2	-7.71	116.44	120.30
1	B	713	ARG	NE-CZ-NH1	7.52	124.06	120.30
1	A	268	ASP	CB-CG-OD1	7.29	124.86	118.30
1	A	268	ASP	CB-CG-OD2	-7.17	111.85	118.30
1	B	40	ARG	NE-CZ-NH1	7.04	123.82	120.30
1	A	132	ARG	NE-CZ-NH1	6.98	123.79	120.30
1	B	649[A]	ARG	NE-CZ-NH2	-6.94	116.83	120.30
1	B	649[B]	ARG	NE-CZ-NH2	-6.94	116.83	120.30
1	A	688	ARG	NE-CZ-NH1	6.71	123.65	120.30
1	A	149	ARG	NE-CZ-NH1	6.66	123.63	120.30
1	B	82	ARG	NE-CZ-NH2	-6.59	117.00	120.30
1	A	730	ARG	NE-CZ-NH1	6.58	123.59	120.30
1	A	504	ARG	NE-CZ-NH2	-6.49	117.06	120.30
1	B	730[A]	ARG	CB-CA-C	6.31	123.02	110.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	730[B]	ARG	CB-CA-C	6.31	123.02	110.40
1	B	199	ASP	CB-CG-OD1	-6.29	112.64	118.30
1	A	504	ARG	NE-CZ-NH1	6.28	123.44	120.30
1	B	497	ARG	NE-CZ-NH2	-6.18	117.21	120.30
1	B	93	ASP	CB-CG-OD1	5.96	123.67	118.30
1	A	314	ARG	NE-CZ-NH2	5.93	123.26	120.30
1	B	393	ARG	NE-CZ-NH1	5.92	123.26	120.30
1	B	426	ARG	NE-CZ-NH1	-5.89	117.35	120.30
1	B	268	ASP	CB-CG-OD2	-5.87	113.01	118.30
1	B	376	ASP	CB-CG-OD2	-5.85	113.03	118.30
1	A	40	ARG	NE-CZ-NH1	5.83	123.22	120.30
1	A	70	GLN	CA-CB-CG	5.77	126.09	113.40
1	B	76	ASP	CB-CG-OD2	-5.76	113.11	118.30
1	B	412	PHE	CB-CG-CD2	-5.58	116.90	120.80
1	A	512	ASP	CB-CG-OD2	-5.56	113.30	118.30
1	A	149	ARG	NE-CZ-NH2	-5.51	117.55	120.30
1	B	610	ARG	NE-CZ-NH1	5.48	123.04	120.30
1	A	587	ASP	CB-CG-OD1	5.47	123.23	118.30
1	B	393	ARG	NE-CZ-NH2	-5.43	117.58	120.30
1	B	55	HIS	CB-CA-C	5.41	121.22	110.40
1	B	268	ASP	CB-CG-OD1	5.41	123.16	118.30
1	A	713	ARG	NE-CZ-NH2	-5.38	117.61	120.30
1	B	504	ARG	NE-CZ-NH1	5.38	122.99	120.30
1	B	76	ASP	CB-CG-OD1	5.35	123.11	118.30
1	A	199	ASP	CB-CG-OD1	-5.33	113.50	118.30
1	B	587	ASP	CB-CG-OD2	-5.33	113.50	118.30
1	B	683	ASP	CB-CG-OD2	-5.32	113.52	118.30
1	A	746	ASP	CB-CG-OD1	5.30	123.07	118.30
1	A	459	ASP	CB-CG-OD2	-5.27	113.56	118.30
1	A	438	GLU	OE1-CD-OE2	5.24	129.59	123.30
1	A	434	TYR	CB-CG-CD1	5.23	124.14	121.00
1	B	540	ARG	NE-CZ-NH1	5.22	122.91	120.30
1	A	610	ARG	NE-CZ-NH1	5.15	122.87	120.30
1	A	146	ASP	CB-CG-OD1	5.14	122.93	118.30
1	A	255[A]	ARG	CB-CG-CD	5.13	124.94	111.60
1	A	255[B]	ARG	CB-CG-CD	5.13	124.94	111.60
1	B	146	ASP	CB-CG-OD1	5.13	122.92	118.30
1	A	403	ARG	NE-CZ-NH1	5.04	122.82	120.30

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	5534	0	5348	13	0
1	B	5544	0	5369	26	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	1	0	0	0	0
4	B	1	0	0	0	0
5	A	2	0	0	0	0
5	B	2	0	0	0	0
6	A	5	0	0	0	0
6	B	5	0	0	0	0
7	A	16	0	28	1	0
7	B	24	0	42	13	0
8	A	743	0	0	6	0
8	B	715	0	0	10	0
All	All	12680	0	10847	49	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 (49) 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)
7:B:808:MPD:O4	7:B:808:MPD:H11	1.50	1.09
1:B:730[A]:ARG:HD2	8:B:1508:HOH:O	1.70	0.92
7:B:808:MPD:O4	7:B:808:MPD:C1	2.15	0.89
1:B:506:ARG:HE	7:B:806:MPD:H12	1.39	0.86
1:A:119[B]:THR:HG21	8:A:982:HOH:O	1.78	0.84
1:A:119[B]:THR:HG23	1:A:593:VAL:HG11	1.61	0.81
7:B:806:MPD:H51	8:B:1258:HOH:O	1.86	0.74
1:B:55:HIS:HD2	8:B:1278:HOH:O	1.72	0.71
1:B:76:ASP:OD2	8:B:901:HOH:O	2.08	0.70
1:B:512:ASP:OD1	8:B:902:HOH:O	2.10	0.69
7:B:806:MPD:H52	7:B:806:MPD:H11	1.75	0.69

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:119[B]:THR:HG23	8:B:1064:HOH:O	1.93	0.68
1:B:629:PRO:O	1:B:633[A]:VAL:HG23	1.96	0.66
1:B:540:ARG:CZ	1:B:540:ARG:HA	2.27	0.64
7:B:806:MPD:H52	7:B:806:MPD:C1	2.28	0.63
1:A:694:GLU:HG2	8:A:903:HOH:O	1.99	0.62
1:B:119[B]:THR:HG22	1:B:593:VAL:HG21	1.83	0.60
1:B:512:ASP:HB2	8:B:1090:HOH:O	2.01	0.60
1:B:290:ALA:HB1	7:B:808:MPD:HM3	1.84	0.59
1:B:520[B]:GLN:OE1	1:B:647:GLN:OE1	2.27	0.53
1:B:662:ASN:H	1:B:725:GLN:HE22	1.56	0.53
1:B:431:ARG:HD2	1:B:447:GLN:OE1	2.09	0.52
1:B:633[A]:VAL:HG22	1:B:719:TYR:CE2	2.45	0.52
1:B:399:GLU:OE1	1:B:403:ARG:NH1	2.43	0.51
1:A:343:GLU:HG2	8:A:1441:HOH:O	2.10	0.51
1:A:431:ARG:HD2	1:A:447:GLN:OE1	2.11	0.51
1:A:119[A]:THR:CG2	1:A:265:ALA:HB2	2.41	0.51
1:B:540:ARG:HA	1:B:540:ARG:NH1	2.25	0.50
7:B:807:MPD:O4	7:B:807:MPD:CM	2.60	0.49
7:B:806:MPD:H13	8:B:1519:HOH:O	2.11	0.49
1:B:119[B]:THR:HG21	8:B:1532:HOH:O	2.10	0.49
1:B:506:ARG:NE	7:B:806:MPD:H12	2.20	0.49
1:A:51:ILE:HD11	1:A:618:VAL:HG12	1.93	0.49
7:B:807:MPD:O4	7:B:807:MPD:HM2	2.16	0.46
7:A:806:MPD:H53	7:A:806:MPD:H11	1.98	0.46
1:A:428:MET:O	1:A:433:ARG:HD3	2.16	0.45
1:A:696:LYS:NZ	8:A:924:HOH:O	2.49	0.45
1:B:591:MET:SD	1:B:594:LEU:HD12	2.57	0.45
1:B:506:ARG:HE	7:B:806:MPD:C1	2.22	0.44
1:A:454:ASP:OD1	1:A:454:ASP:N	2.51	0.43
1:B:649[A]:ARG:HD3	1:B:649[A]:ARG:HA	1.70	0.43
1:B:55:HIS:CD2	8:B:1278:HOH:O	2.59	0.43
7:B:808:MPD:H11	7:B:808:MPD:HO4	1.74	0.42
1:B:428:MET:O	1:B:433:ARG:HD3	2.19	0.42
1:A:255[A]:ARG:HG3	8:A:1300:HOH:O	2.19	0.42
1:B:633[A]:VAL:HG22	1:B:719:TYR:CZ	2.55	0.41
1:B:454:ASP:OD1	1:B:454:ASP:N	2.51	0.41
1:A:144:ASN:HA	1:A:146:ASP:OD1	2.20	0.41
1:A:426[B]:ARG:NH2	8:A:907:HOH:O	2.54	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	714/728 (98%)	704 (99%)	10 (1%)	0	100	100
1	B	717/728 (98%)	706 (98%)	11 (2%)	0	100	100
All	All	1431/1456 (98%)	1410 (98%)	21 (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	553/560 (99%)	544 (98%)	9 (2%)	62	49
1	B	556/560 (99%)	545 (98%)	11 (2%)	55	40
All	All	1109/1120 (99%)	1089 (98%)	20 (2%)	62	45

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	40	ARG
1	A	64	LYS
1	A	73	GLU
1	A	118	ARG
1	A	141	ASP
1	A	608	LYS
1	A	610	ARG
1	A	694	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	725	GLN
1	B	40	ARG
1	B	64	LYS
1	B	74	LYS
1	B	118	ARG
1	B	141	ASP
1	B	610	ARG
1	B	633[A]	VAL
1	B	633[B]	VAL
1	B	647	GLN
1	B	649[A]	ARG
1	B	649[B]	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	70	GLN
1	A	227	ASN
1	A	247	ASN
1	A	339	HIS
1	A	544	GLN
1	A	647	GLN
1	B	46	GLN
1	B	55	HIS
1	B	568	HIS
1	B	647	GLN
1	B	650	HIS
1	B	725	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)

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	B	111[B]	-	10,17,18	1.99	4 (40%)	10,23,25	2.06	2 (20%)
1	TOX	A	111[B]	-	10,17,18	2.12	4 (40%)	10,23,25	2.62	5 (50%)
1	TOX	A	111[A]	2	10,17,18	2.12	4 (40%)	10,23,25	2.62	5 (50%)
1	TOX	B	111[A]	2	10,17,18	1.99	4 (40%)	10,23,25	2.06	2 (20%)

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	B	111[B]	-	-	2/4/8/10	0/2/2/2
1	TOX	A	111[B]	-	-	2/4/8/10	0/2/2/2
1	TOX	A	111[A]	2	-	2/4/8/10	0/2/2/2
1	TOX	B	111[A]	2	-	2/4/8/10	0/2/2/2

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	111[A]	TOX	O-C	3.50	1.33	1.19
1	A	111[B]	TOX	O-C	3.50	1.33	1.19
1	B	111[A]	TOX	CD1-NE1	-3.26	1.36	1.39
1	B	111[B]	TOX	CD1-NE1	-3.26	1.36	1.39
1	A	111[A]	TOX	CE3-CD2	-3.08	1.35	1.42
1	A	111[B]	TOX	CE3-CD2	-3.08	1.35	1.42
1	B	111[A]	TOX	CE3-CD2	-3.06	1.36	1.42
1	B	111[B]	TOX	CE3-CD2	-3.06	1.36	1.42
1	A	111[A]	TOX	CD1-NE1	-2.99	1.36	1.39
1	A	111[B]	TOX	CD1-NE1	-2.99	1.36	1.39
1	B	111[A]	TOX	O-C	2.80	1.31	1.19
1	B	111[B]	TOX	O-C	2.80	1.31	1.19
1	A	111[A]	TOX	CH2-CZ3	2.53	1.44	1.38
1	A	111[B]	TOX	CH2-CZ3	2.53	1.44	1.38
1	B	111[A]	TOX	CH2-CZ3	2.30	1.44	1.38
1	B	111[B]	TOX	CH2-CZ3	2.30	1.44	1.38

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	111[A]	TOX	CB-CG-CD1	-4.77	122.08	127.97
1	B	111[B]	TOX	CB-CG-CD1	-4.77	122.08	127.97
1	A	111[A]	TOX	CB-CG-CD1	-4.63	122.24	127.97
1	A	111[B]	TOX	CB-CG-CD1	-4.63	122.24	127.97
1	A	111[A]	TOX	CZ3-CE3-CD2	-4.00	115.35	120.89
1	A	111[B]	TOX	CZ3-CE3-CD2	-4.00	115.35	120.89
1	A	111[A]	TOX	CB-CG-CD2	3.47	131.65	126.25
1	A	111[B]	TOX	CB-CG-CD2	3.47	131.65	126.25
1	A	111[A]	TOX	CE3-CD2-CG	-3.18	128.58	134.42
1	A	111[B]	TOX	CE3-CD2-CG	-3.18	128.58	134.42
1	B	111[A]	TOX	CB-CG-CD2	2.72	130.49	126.25
1	B	111[B]	TOX	CB-CG-CD2	2.72	130.49	126.25
1	A	111[A]	TOX	CB-CA-C	2.26	115.70	111.47
1	A	111[B]	TOX	CB-CA-C	2.26	115.70	111.47

There are no chirality outliers.

All (8) torsion outliers are listed below:

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

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 15 ligands modelled in this entry, 4 are monoatomic - leaving 11 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	OXY	A	804	-	1,1,1	0.25	0	-		
6	PO4	B	805	-	4,4,4	1.12	0	6,6,6	1.75	2 (33%)
7	MPD	A	807	-	7,7,7	0.59	0	9,10,10	1.51	2 (22%)
7	MPD	B	807	-	7,7,7	0.62	0	9,10,10	1.86	2 (22%)
7	MPD	B	806	-	7,7,7	1.48	1 (14%)	9,10,10	1.58	3 (33%)
7	MPD	B	808	-	7,7,7	0.91	0	9,10,10	1.64	3 (33%)
7	MPD	A	806	-	7,7,7	0.88	0	9,10,10	1.76	3 (33%)
2	HEM	B	801	1	41,50,50	1.27	6 (14%)	45,82,82	1.66	11 (24%)
6	PO4	A	805	-	4,4,4	1.17	0	6,6,6	3.98	3 (50%)
5	OXY	B	804	-	1,1,1	0.35	0	-		
2	HEM	A	801	1	41,50,50	1.57	5 (12%)	45,82,82	2.11	21 (46%)

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
7	MPD	A	807	-	-	1/5/5/5	-
7	MPD	B	807	-	-	4/5/5/5	-
7	MPD	B	806	-	-	3/5/5/5	-
7	MPD	B	808	-	-	2/5/5/5	-
7	MPD	A	806	-	-	0/5/5/5	-
2	HEM	B	801	1	-	2/12/54/54	-
2	HEM	A	801	1	-	2/12/54/54	-

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	HEM	C4D-ND	-6.13	1.29	1.40
2	B	801	HEM	C1B-NB	-3.91	1.33	1.40
2	A	801	HEM	C4B-NB	-3.75	1.31	1.38
7	B	806	MPD	C3-C2	3.12	1.62	1.53

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	801	HEM	C4D-ND	-2.67	1.35	1.40
2	B	801	HEM	C4A-NA	2.60	1.41	1.36
2	A	801	HEM	C1B-NB	-2.46	1.36	1.40
2	B	801	HEM	CHB-C1B	2.38	1.41	1.35
2	A	801	HEM	CMD-C2D	-2.19	1.46	1.50
2	A	801	HEM	O1D-CGD	2.16	1.29	1.22
2	B	801	HEM	O2A-CGA	-2.16	1.23	1.30
2	B	801	HEM	CHC-C4B	-2.01	1.35	1.41

All (50) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	805	PO4	O4-P-O3	6.87	130.04	107.97
6	A	805	PO4	O4-P-O1	-6.36	87.63	110.89
7	B	807	MPD	O4-C4-C5	-4.62	89.36	109.38
2	B	801	HEM	CBA-CAA-C2A	4.52	120.33	112.62
2	A	801	HEM	C2C-C3C-C4C	4.45	110.00	106.90
2	A	801	HEM	CHB-C1B-NB	3.83	129.12	124.38
2	A	801	HEM	C4C-CHD-C1D	3.78	127.55	122.56
2	B	801	HEM	CHA-C4D-ND	3.61	128.84	124.38
2	A	801	HEM	CHC-C4B-NB	3.59	128.33	124.43
2	A	801	HEM	O2A-CGA-CBA	3.54	125.39	114.03
2	A	801	HEM	CMD-C2D-C1D	3.24	129.97	125.04
7	A	807	MPD	CM-C2-C1	3.13	117.10	110.57
2	A	801	HEM	C3D-C4D-ND	2.96	113.46	110.17
2	A	801	HEM	CAA-CBA-CGA	-2.96	105.46	113.76
2	B	801	HEM	O2D-CGD-CBD	2.93	123.46	114.03
7	A	806	MPD	O2-C2-CM	-2.89	98.80	108.08
2	B	801	HEM	C4B-CHC-C1C	2.76	126.19	122.56
2	A	801	HEM	C3C-C4C-NC	-2.66	105.92	110.94
2	B	801	HEM	CAD-C3D-C4D	2.62	129.23	124.66
7	A	807	MPD	O2-C2-CM	-2.58	99.79	108.08
2	A	801	HEM	CMC-C2C-C3C	2.55	129.44	124.68
2	A	801	HEM	C1D-C2D-C3D	-2.54	104.28	106.96
2	A	801	HEM	O1A-CGA-CBA	-2.52	114.97	123.08
2	B	801	HEM	CHD-C1D-ND	2.45	127.10	124.43
2	B	801	HEM	CBB-CAB-C3B	-2.42	115.56	127.62
7	B	808	MPD	C1-C2-C3	2.41	121.18	109.96
2	A	801	HEM	C2D-C1D-ND	2.41	112.77	109.88
6	B	805	PO4	O2-P-O1	-2.41	102.08	110.89
7	B	806	MPD	O2-C2-C1	-2.40	100.39	108.08
7	B	806	MPD	O4-C4-C3	2.39	121.00	111.36

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	801	HEM	CAD-C3D-C2D	-2.39	123.43	127.88
2	A	801	HEM	C3B-C2B-C1B	-2.38	104.72	106.49
6	A	805	PO4	O2-P-O1	2.36	119.55	110.89
7	B	807	MPD	CM-C2-C1	-2.33	105.71	110.57
2	A	801	HEM	CBD-CAD-C3D	2.32	119.07	112.63
7	A	806	MPD	CM-C2-C3	2.28	120.56	109.96
2	A	801	HEM	CHB-C1B-C2B	-2.26	120.48	126.72
2	B	801	HEM	CHA-C4D-C3D	-2.25	121.10	125.33
2	A	801	HEM	C1B-NB-C4B	2.25	107.39	105.07
2	A	801	HEM	CBA-CAA-C2A	2.24	116.44	112.62
2	A	801	HEM	C4A-C3A-C2A	-2.20	105.46	107.00
2	A	801	HEM	O1D-CGD-CBD	-2.20	116.02	123.08
2	B	801	HEM	O1A-CGA-CBA	-2.19	116.05	123.08
6	B	805	PO4	O3-P-O1	2.17	118.85	110.89
7	B	806	MPD	C1-C2-C3	2.17	120.06	109.96
7	A	806	MPD	O2-C2-C3	-2.16	101.68	109.80
7	B	808	MPD	O4-C4-C5	-2.12	100.19	109.38
7	B	808	MPD	CM-C2-C1	2.08	114.90	110.57
2	A	801	HEM	CAD-C3D-C4D	2.07	128.28	124.66
2	B	801	HEM	C4B-C3B-C2B	-2.04	105.49	107.11

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	806	MPD	C1-C2-C3-C4
7	B	806	MPD	O2-C2-C3-C4
7	B	807	MPD	CM-C2-C3-C4
7	B	807	MPD	C2-C3-C4-C5
7	B	808	MPD	C2-C3-C4-C5
7	B	808	MPD	C2-C3-C4-O4
2	A	801	HEM	CAA-CBA-CGA-O1A
2	A	801	HEM	CAA-CBA-CGA-O2A
2	B	801	HEM	CAA-CBA-CGA-O2A
7	A	807	MPD	O2-C2-C3-C4
7	B	807	MPD	O2-C2-C3-C4
2	B	801	HEM	CAA-CBA-CGA-O1A
7	B	806	MPD	C2-C3-C4-C5
7	B	807	MPD	C2-C3-C4-O4

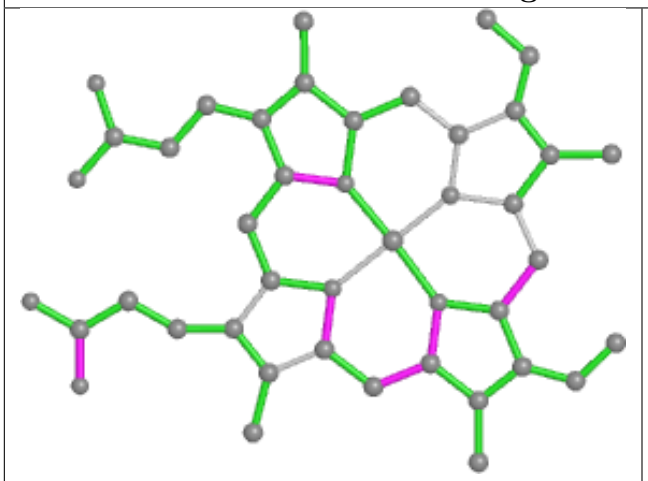
There are no ring outliers.

4 monomers are involved in 14 short contacts:

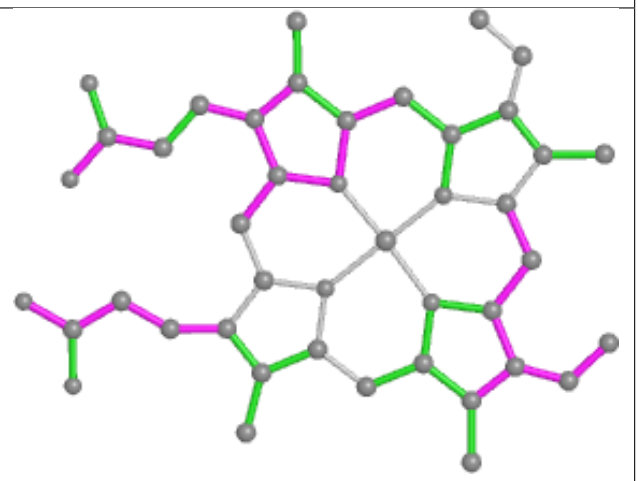
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	807	MPD	2	0
7	B	806	MPD	7	0
7	B	808	MPD	4	0
7	A	806	MPD	1	0

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.

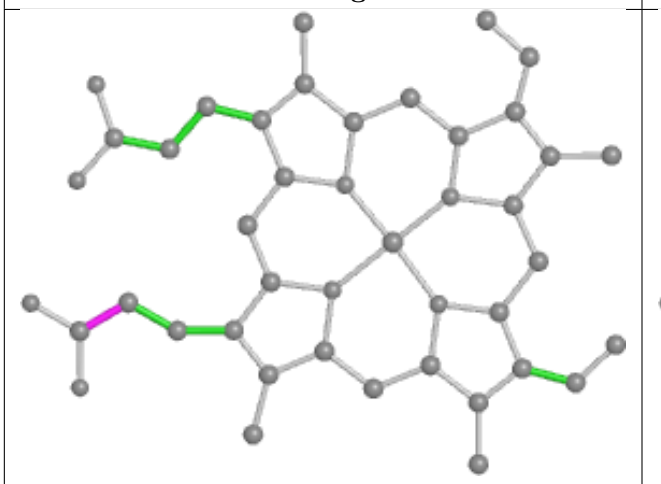
Ligand HEM B 801



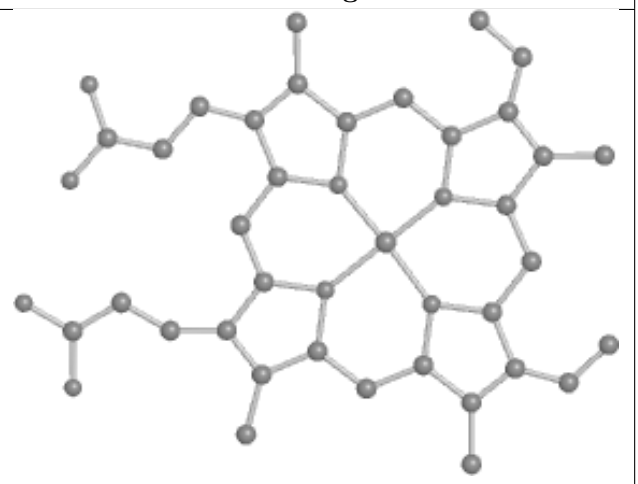
Bond lengths



Bond angles

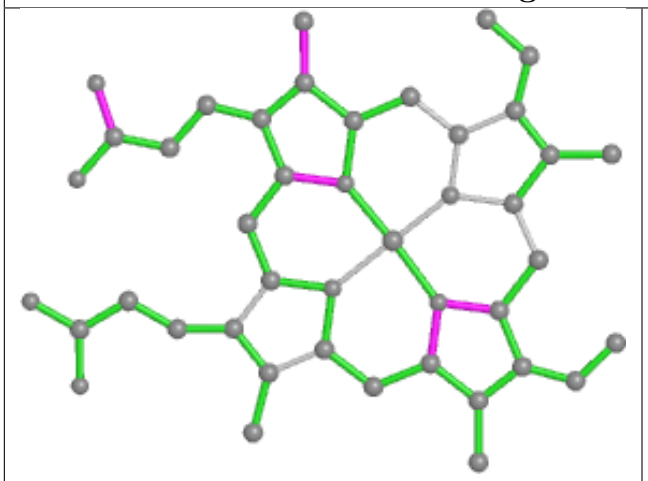


Torsions

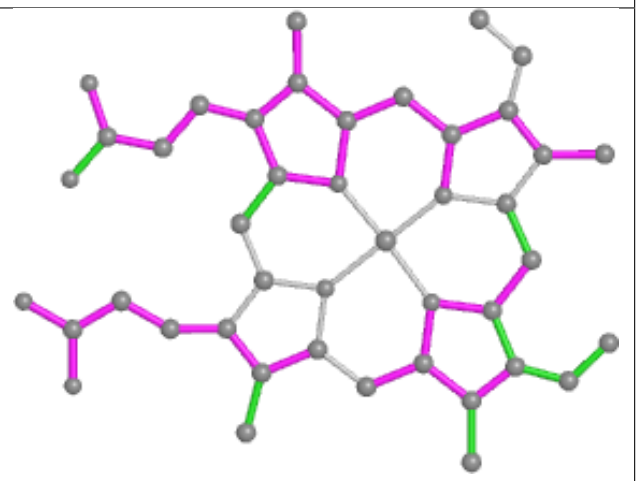


Rings

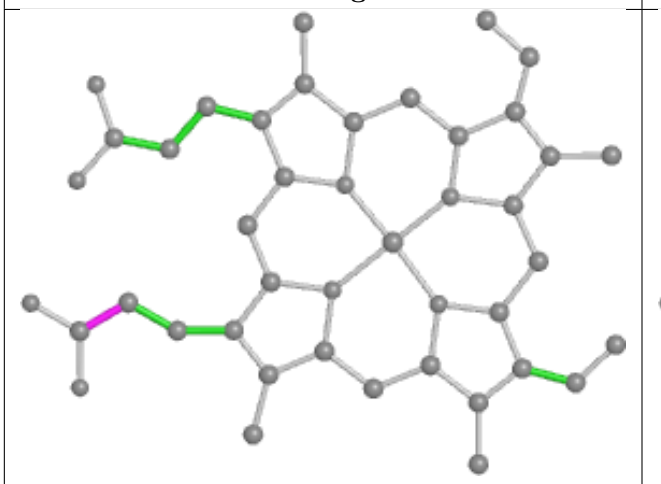
Ligand HEM A 801



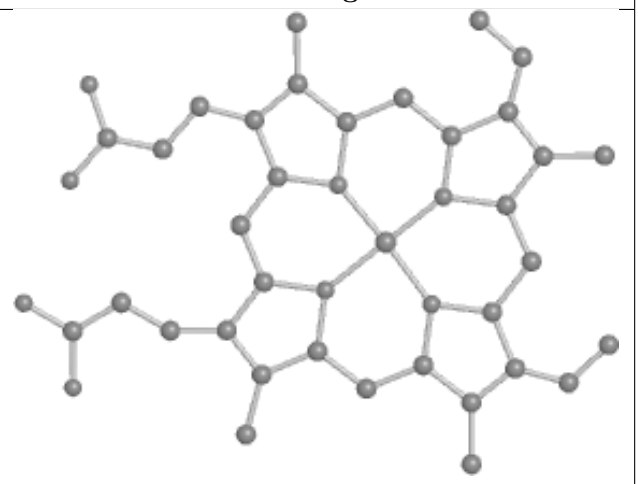
Bond lengths



Bond angles



Torsions



Rings

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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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%)	-0.50	10 (1%)	75	76	16, 23, 42, 85	0
1	B	712/728 (97%)	-0.49	7 (0%)	82	82	16, 23, 41, 86	0
All	All	1424/1456 (97%)	-0.50	17 (1%)	79	79	16, 23, 42, 86	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	748	ALA	4.6
1	A	748	ALA	4.2
1	B	610	ARG	4.1
1	A	540	ARG	4.1
1	B	540	ARG	4.0
1	B	679	ALA	3.7
1	A	608	LYS	2.9
1	A	680	ALA	2.9
1	A	679	ALA	2.8
1	A	221	GLY	2.4
1	A	454	ASP	2.4
1	A	213	GLY	2.4
1	B	541	GLY	2.4
1	B	680	ALA	2.1
1	B	454	ASP	2.1
1	A	219	TYR	2.0
1	A	215	PRO	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, 95th 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	B	111[A]	16/17	0.96	0.15	15,18,25,28	0
1	TOX	B	111[B]	16/17	0.96	0.15	15,18,20,25	1
1	TOX	A	111[A]	16/17	0.98	0.12	15,17,21,21	2
1	TOX	A	111[B]	16/17	0.98	0.12	15,17,19,21	2

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

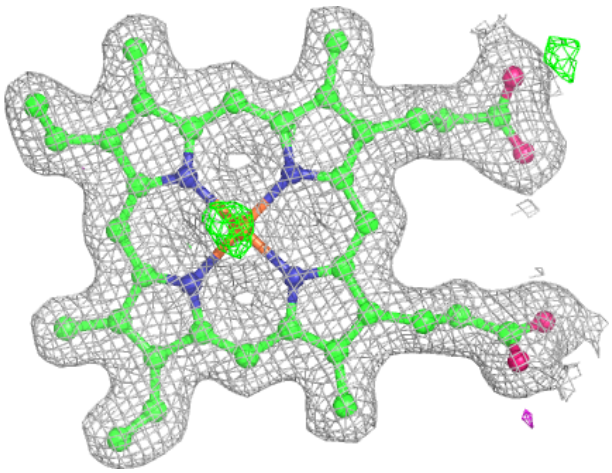
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

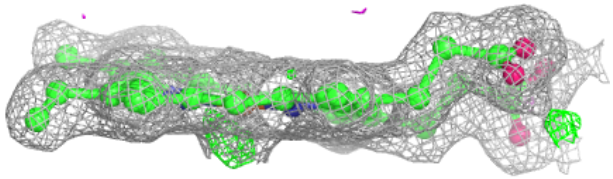
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	PO4	A	805	5/5	0.86	0.18	48,54,68,72	0
7	MPD	B	808	8/8	0.86	0.25	57,66,75,78	0
7	MPD	B	806	8/8	0.88	0.17	41,50,52,54	0
7	MPD	A	807	8/8	0.91	0.17	59,61,67,69	0
7	MPD	B	807	8/8	0.93	0.16	35,40,55,62	0
5	OXY	B	804	2/2	0.94	0.12	34,34,34,37	0
7	MPD	A	806	8/8	0.94	0.12	39,41,55,58	0
5	OXY	A	804	2/2	0.95	0.21	36,36,36,45	0
6	PO4	B	805	5/5	0.96	0.18	37,51,57,60	0
4	CL	A	803	1/1	0.96	0.06	39,39,39,39	0
4	CL	B	803	1/1	0.97	0.06	39,39,39,39	0
2	HEM	B	801	43/43	0.99	0.11	15,17,19,23	0
2	HEM	A	801	43/43	0.99	0.09	15,17,20,22	0
3	NA	B	802	1/1	1.00	0.02	20,20,20,20	0
3	NA	A	802	1/1	1.00	0.03	19,19,19,19	0

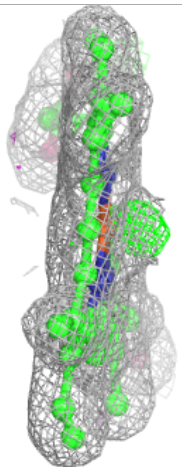
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

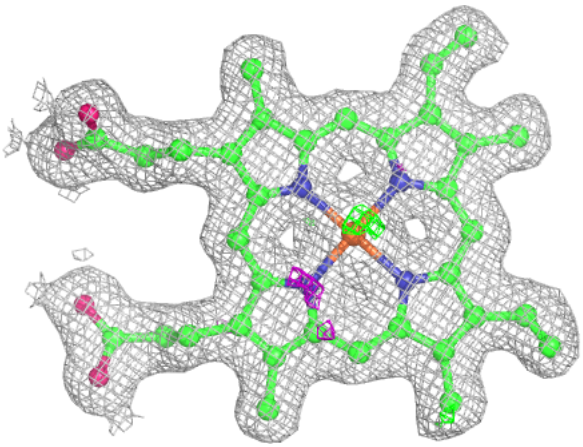
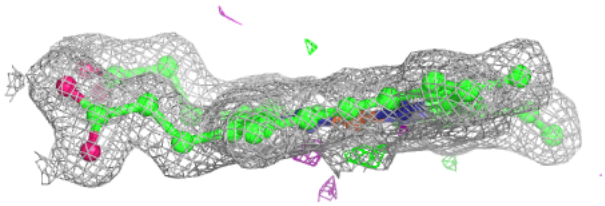
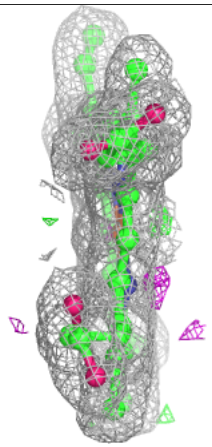
Electron density around HEM B 801:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)







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)	
	

6.5 Other polymers (i)

There are no such residues in this entry.