



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 6, 2023 – 10:09 PM EST

PDB ID : 5SX2  
Title : Crystal structure of the D141E mutant of *B. pseudomallei* KatG at pH 8.0.  
Authors : Loewen, P.C.  
Deposited on : 2016-08-09  
Resolution : 2.15 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (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

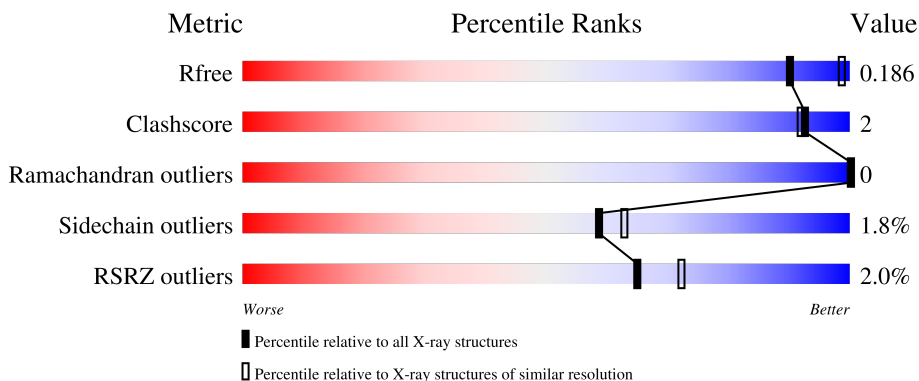
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

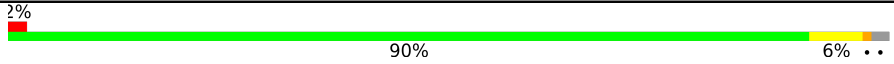
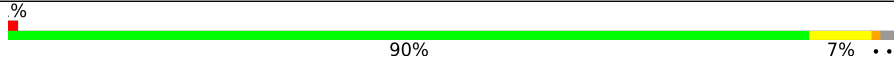
The reported resolution of this entry is 2.15 Å.

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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	728	 2% 90% 6% ..
1	B	728	 % 90% 7% ..

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
4	OXY	A	804	-	-	X	-

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 12471 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	Total	C	N	O	S	0	1	0
			5509	3478	980	1037	14			
1	B	713	Total	C	N	O	S	0	2	0
			5516	3482	983	1037	14			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	141	GLU	ASP	engineered mutation	UNP Q3JNW6
B	141	GLU	ASP	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	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

*Continued on next page...*

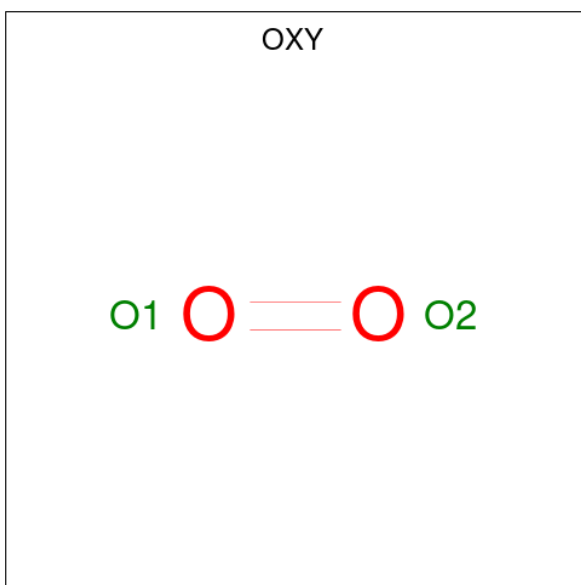
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			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 OXYGEN MOLECULE (three-letter code: OXY) (formula: O<sub>2</sub>).



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

- Molecule 5 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



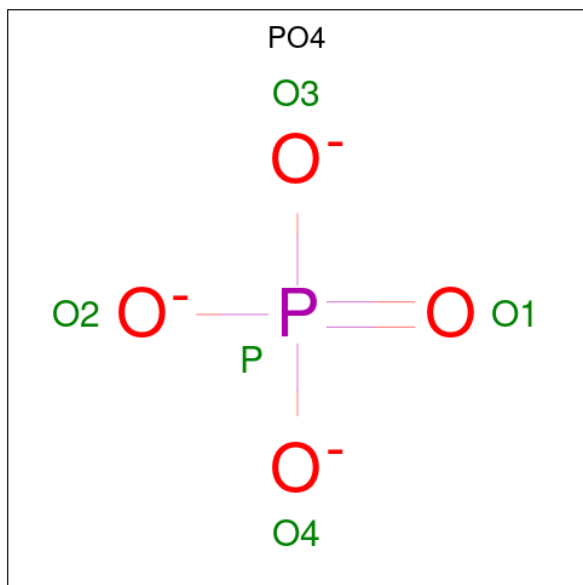
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 8 6 2	0	0
5	A	1	Total C O 8 6 2	0	0
5	B	1	Total C O 8 6 2	0	0

- Molecule 6 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula:  $C_4H_{12}NO_3$ ).



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

- Molecule 7 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



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

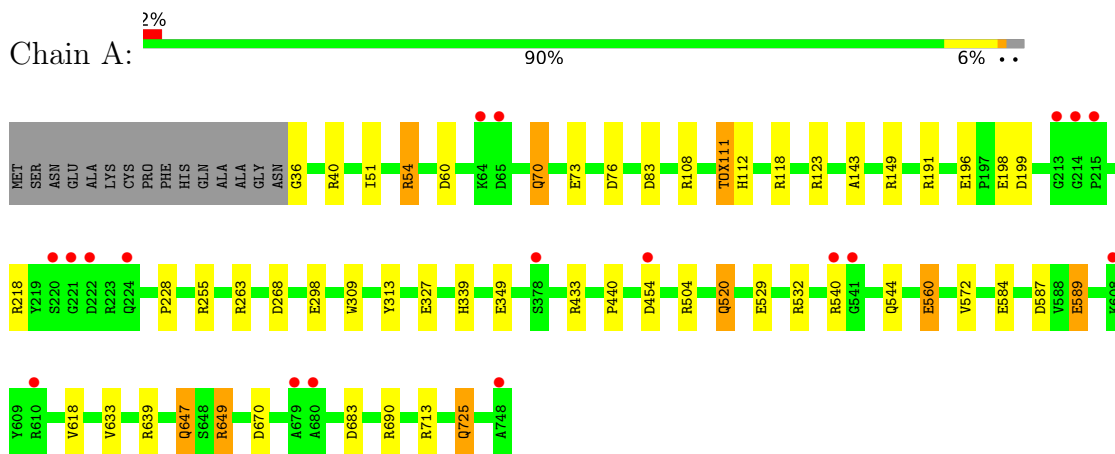
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	624	Total	O	0	0
			624	624		
8	B	676	Total	O	0	0
			676	676		

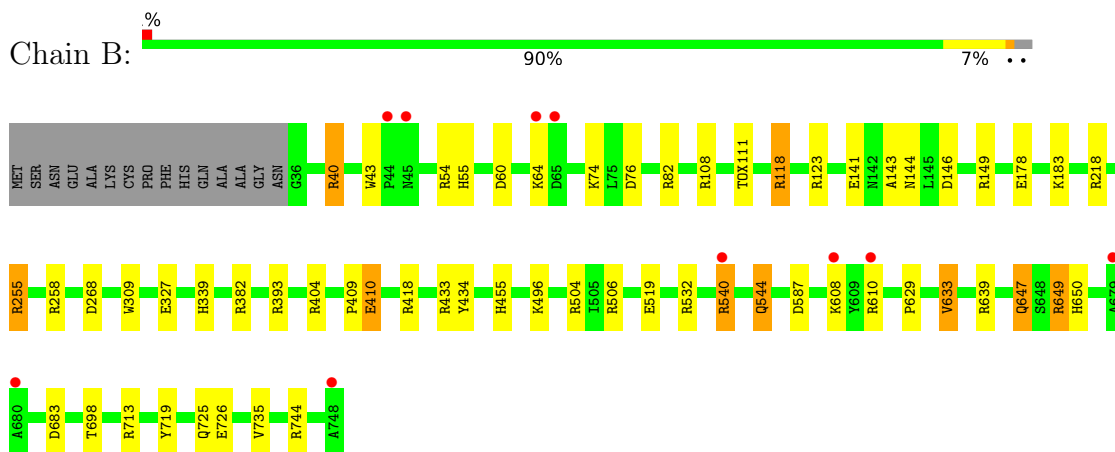
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Catalase-peroxidase



- Molecule 1: Catalase-peroxidase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.33Å 114.82Å 174.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.15 29.05 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.0 (20.00-2.15) 99.1 (29.05-2.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.14 (at 2.16Å)	Xtrriage
Refinement program	REFMAC 5.8.0151	Depositor
R, $R_{free}$	0.142 , 0.180 0.153 , 0.186	Depositor DCC
$R_{free}$ test set	5372 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.6	Xtrriage
Anisotropy	0.496	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 41.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12471	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.82% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, PO4, OXY, MPD, TOX, TRS, HEM

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	1.18	12/5637 (0.2%)	1.13	40/7661 (0.5%)
1	B	1.19	8/5648 (0.1%)	1.16	51/7675 (0.7%)
All	All	1.18	20/11285 (0.2%)	1.14	91/15336 (0.6%)

All (20) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	410	GLU	CD-OE1	11.59	1.38	1.25
1	B	726	GLU	CG-CD	8.95	1.65	1.51
1	B	532	ARG	CD-NE	-7.50	1.33	1.46
1	B	410	GLU	CD-OE2	6.66	1.32	1.25
1	A	647	GLN	CG-CD	6.24	1.65	1.51
1	A	532	ARG	CD-NE	-6.02	1.36	1.46
1	A	327	GLU	CD-OE1	5.99	1.32	1.25
1	A	725	GLN	CD-NE2	5.92	1.47	1.32
1	A	349	GLU	CD-OE2	-5.80	1.19	1.25
1	A	73	GLU	CG-CD	5.74	1.60	1.51
1	A	313	TYR	CE2-CZ	5.67	1.46	1.38
1	A	584	GLU	CG-CD	5.67	1.60	1.51
1	A	36	GLY	N-CA	5.67	1.54	1.46
1	B	410	GLU	CG-CD	5.66	1.60	1.51
1	B	393	ARG	CZ-NH2	-5.62	1.25	1.33
1	B	327	GLU	CD-OE1	5.47	1.31	1.25
1	A	560	GLU	CD-OE2	-5.21	1.20	1.25
1	B	118	ARG	CZ-NH1	5.16	1.39	1.33
1	A	198	GLU	CG-CD	5.11	1.59	1.51
1	A	589	GLU	CD-OE2	5.00	1.31	1.25

All (91) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	532	ARG	NE-CZ-NH2	-22.81	108.89	120.30
1	B	532	ARG	NE-CZ-NH2	-19.89	110.36	120.30
1	A	54	ARG	NE-CZ-NH2	-12.17	114.21	120.30
1	B	532	ARG	NE-CZ-NH1	12.15	126.37	120.30
1	A	454	ASP	CB-CG-OD2	12.10	129.19	118.30
1	A	532	ARG	NE-CZ-NH1	12.09	126.35	120.30
1	B	108	ARG	NE-CZ-NH2	-11.09	114.76	120.30
1	B	54	ARG	NE-CZ-NH2	-10.55	115.02	120.30
1	A	649	ARG	NE-CZ-NH1	10.30	125.45	120.30
1	A	76	ASP	CB-CG-OD2	-9.95	109.35	118.30
1	B	54	ARG	CG-CD-NE	-9.18	92.53	111.80
1	B	393	ARG	NE-CZ-NH2	-8.38	116.11	120.30
1	B	82	ARG	NE-CZ-NH2	-8.22	116.19	120.30
1	A	54	ARG	NE-CZ-NH1	8.08	124.34	120.30
1	A	149	ARG	NE-CZ-NH1	8.01	124.30	120.30
1	A	60	ASP	CB-CG-OD1	7.85	125.37	118.30
1	B	82	ARG	NE-CZ-NH1	7.78	124.19	120.30
1	A	263	ARG	NE-CZ-NH2	-7.63	116.48	120.30
1	B	393	ARG	NE-CZ-NH1	7.63	124.12	120.30
1	B	54	ARG	NE-CZ-NH1	7.52	124.06	120.30
1	A	218	ARG	NE-CZ-NH2	-7.32	116.64	120.30
1	B	76	ASP	CB-CG-OD2	-7.27	111.75	118.30
1	A	108	ARG	NE-CZ-NH1	7.18	123.89	120.30
1	A	670	ASP	CB-CG-OD2	-7.17	111.85	118.30
1	A	670	ASP	CB-CG-OD1	7.13	124.72	118.30
1	B	713	ARG	NE-CZ-NH1	7.08	123.84	120.30
1	B	713	ARG	NE-CZ-NH2	-7.01	116.79	120.30
1	B	108	ARG	NE-CZ-NH1	7.00	123.80	120.30
1	A	123	ARG	NE-CZ-NH1	6.96	123.78	120.30
1	B	506	ARG	NE-CZ-NH1	6.93	123.76	120.30
1	B	255[A]	ARG	NE-CZ-NH1	6.87	123.74	120.30
1	B	255[B]	ARG	NE-CZ-NH1	6.87	123.74	120.30
1	B	649	ARG	NE-CZ-NH1	6.76	123.68	120.30
1	B	382	ARG	NE-CZ-NH2	-6.73	116.94	120.30
1	A	268	ASP	CB-CG-OD2	-6.62	112.35	118.30
1	A	713	ARG	NE-CZ-NH1	6.56	123.58	120.30
1	B	60	ASP	CB-CG-OD2	-6.56	112.39	118.30
1	B	123	ARG	NE-CZ-NH1	6.50	123.55	120.30
1	B	60	ASP	CB-CG-OD1	6.48	124.13	118.30
1	B	639	ARG	NE-CZ-NH1	6.48	123.54	120.30
1	A	532	ARG	CD-NE-CZ	6.45	132.63	123.60
1	B	587	ASP	CB-CG-OD1	6.41	124.07	118.30
1	A	454	ASP	CB-CG-OD1	-6.38	112.56	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	587	ASP	CB-CG-OD1	6.37	124.03	118.30
1	B	683	ASP	CB-CG-OD1	6.36	124.02	118.30
1	B	649	ARG	NE-CZ-NH2	-6.33	117.14	120.30
1	A	540	ARG	NE-CZ-NH1	6.32	123.46	120.30
1	A	255	ARG	NE-CZ-NH1	6.31	123.46	120.30
1	B	683	ASP	CB-CG-OD2	-6.25	112.67	118.30
1	A	532	ARG	CG-CD-NE	-6.24	98.69	111.80
1	A	504	ARG	NE-CZ-NH1	6.22	123.41	120.30
1	A	683	ASP	CB-CG-OD1	6.21	123.89	118.30
1	B	610	ARG	NE-CZ-NH1	6.17	123.39	120.30
1	B	532	ARG	CD-NE-CZ	6.16	132.22	123.60
1	A	60	ASP	CB-CG-OD2	-6.13	112.78	118.30
1	B	433	ARG	NE-CZ-NH2	-6.03	117.28	120.30
1	B	258	ARG	NE-CZ-NH1	6.02	123.31	120.30
1	A	218	ARG	NE-CZ-NH1	6.01	123.31	120.30
1	B	255[A]	ARG	NE-CZ-NH2	-5.99	117.31	120.30
1	B	255[B]	ARG	NE-CZ-NH2	-5.99	117.31	120.30
1	B	218	ARG	NE-CZ-NH2	-5.92	117.34	120.30
1	B	404	ARG	NE-CZ-NH2	-5.84	117.38	120.30
1	B	433	ARG	NE-CZ-NH1	5.83	123.22	120.30
1	A	263	ARG	NE-CZ-NH1	5.83	123.21	120.30
1	A	633	VAL	CG1-CB-CG2	-5.83	101.58	110.90
1	B	532	ARG	CG-CD-NE	-5.82	99.58	111.80
1	B	218	ARG	NE-CZ-NH1	5.82	123.21	120.30
1	B	40	ARG	NE-CZ-NH1	5.76	123.18	120.30
1	A	191	ARG	NE-CZ-NH2	-5.75	117.43	120.30
1	A	433	ARG	NE-CZ-NH1	5.74	123.17	120.30
1	B	149	ARG	NE-CZ-NH1	5.69	123.14	120.30
1	A	70	GLN	CA-CB-CG	5.57	125.66	113.40
1	B	744	ARG	NE-CZ-NH2	-5.53	117.53	120.30
1	B	735	VAL	CG1-CB-CG2	-5.45	102.17	110.90
1	B	540	ARG	NE-CZ-NH1	5.45	123.03	120.30
1	B	519	GLU	CG-CD-OE1	-5.38	107.55	118.30
1	A	83	ASP	CB-CG-OD1	-5.33	113.51	118.30
1	A	639	ARG	NE-CZ-NH1	5.29	122.95	120.30
1	B	633	VAL	CB-CA-C	-5.26	101.40	111.40
1	B	504	ARG	NE-CZ-NH2	-5.25	117.68	120.30
1	A	649	ARG	NE-CZ-NH2	-5.24	117.68	120.30
1	A	529	GLU	OE1-CD-OE2	-5.19	117.08	123.30
1	A	191	ARG	NE-CZ-NH1	5.18	122.89	120.30
1	B	504	ARG	NE-CZ-NH1	5.18	122.89	120.30
1	A	589	GLU	CA-CB-CG	5.16	124.75	113.40

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	123	ARG	NE-CZ-NH2	-5.13	117.74	120.30
1	B	268	ASP	CB-CG-OD2	-5.07	113.74	118.30
1	B	268	ASP	CB-CG-OD1	5.07	122.86	118.30
1	A	196	GLU	OE1-CD-OE2	-5.06	117.23	123.30
1	B	418	ARG	NE-CZ-NH1	5.03	122.81	120.30
1	B	434	TYR	CB-CG-CD1	5.02	124.01	121.00

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5509	0	5319	14	0
1	B	5516	0	5328	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	4	0	0	2	0
4	B	4	0	0	1	0
5	A	16	0	28	3	0
5	B	8	0	14	2	0
6	A	8	0	12	0	0
6	B	8	0	12	0	0
7	A	5	0	0	0	0
7	B	5	0	0	0	0
8	A	624	0	0	7	0
8	B	676	0	0	6	0
All	All	12471	0	10773	33	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 (33) 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)
5:A:805:MPD:CM	5:A:805:MPD:O4	2.38	0.70
1:A:111:TOX:O2	4:A:804:OXY:O1	2.10	0.69
1:A:520:GLN:HG3	8:A:1286:HOH:O	1.92	0.67
1:B:410:GLU:OE1	8:B:901:HOH:O	2.13	0.67
1:A:544:GLN:OE1	8:A:902:HOH:O	2.13	0.66
1:A:725:GLN:OE1	8:A:901:HOH:O	2.13	0.65
5:B:805:MPD:CM	5:B:805:MPD:H52	2.28	0.64
5:B:805:MPD:H52	5:B:805:MPD:HM2	1.80	0.63
1:B:725:GLN:NE2	8:B:903:HOH:O	2.33	0.61
1:A:112:HIS:NE2	4:A:804:OXY:O1	2.36	0.59
1:B:255[B]:ARG:HD3	8:B:1035:HOH:O	2.01	0.59
1:B:650:HIS:HD2	1:B:698:THR:OG1	1.86	0.58
1:B:178:GLU:OE1	8:B:902:HOH:O	2.18	0.54
1:B:55:HIS:HD2	8:B:1213:HOH:O	1.91	0.54
1:A:560:GLU:HG3	1:A:572:VAL:HG23	1.94	0.48
5:A:806:MPD:H53	5:A:806:MPD:H11	1.94	0.48
5:A:805:MPD:O4	5:A:805:MPD:HM1	2.10	0.48
1:A:725:GLN:CD	8:A:901:HOH:O	2.51	0.47
1:B:339:HIS:CD2	1:B:409:PRO:HB3	2.49	0.47
1:B:633:VAL:HG22	1:B:719:TYR:CZ	2.52	0.45
1:B:141:GLU:OE2	4:B:803:OXY:O1	2.33	0.44
1:A:54:ARG:NE	1:A:199:ASP:OD2	2.46	0.43
1:A:228:PRO:HG3	8:A:1201:HOH:O	2.19	0.43
1:B:647:GLN:HG2	8:B:1364:HOH:O	2.18	0.43
1:B:455:HIS:CE1	1:B:544:GLN:HG3	2.55	0.42
1:A:143:ALA:HA	1:A:309:TRP:CZ3	2.55	0.42
1:B:144:ASN:HA	1:B:146:ASP:OD1	2.19	0.42
1:B:143:ALA:HA	1:B:309:TRP:CZ3	2.56	0.41
1:B:629:PRO:O	1:B:633:VAL:HG23	2.20	0.41
1:A:51:ILE:HD11	1:A:618:VAL:HG12	2.01	0.41
1:A:339:HIS:HB2	8:A:1292:HOH:O	2.20	0.41
1:A:298:GLU:HB2	1:B:43:TRP:CE2	2.57	0.40
1:A:725:GLN:CG	8:A:901:HOH:O	2.69	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	711/728 (98%)	698 (98%)	13 (2%)	0	100	100
1	B	712/728 (98%)	700 (98%)	12 (2%)	0	100	100
All	All	1423/1456 (98%)	1398 (98%)	25 (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	550/560 (98%)	541 (98%)	9 (2%)	62	67
1	B	551/560 (98%)	540 (98%)	11 (2%)	55	59
All	All	1101/1120 (98%)	1081 (98%)	20 (2%)	59	63

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

Mol	Chain	Res	Type
1	A	40	ARG
1	A	70	GLN
1	A	118	ARG
1	A	440	PRO
1	A	520	GLN
1	A	589	GLU
1	A	647	GLN
1	A	649	ARG
1	A	690	ARG
1	B	40	ARG
1	B	64	LYS
1	B	74	LYS
1	B	118	ARG
1	B	183	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	496	LYS
1	B	540	ARG
1	B	544	GLN
1	B	608	LYS
1	B	647	GLN
1	B	649	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	70	GLN
1	A	647	GLN
1	B	55	HIS
1	B	339	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](#)

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	2.81	6 (60%)	10,23,25	1.45	1 (10%)
1	TOX	B	111	1	10,17,18	2.75	3 (30%)	10,23,25	1.96	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	1	-	2/4/8/10	0/2/2/2

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	111	TOX	O-C	6.11	1.44	1.19
1	A	111	TOX	O-C	5.16	1.40	1.19
1	B	111	TOX	CD1-NE1	-5.15	1.34	1.39
1	A	111	TOX	CD1-NE1	-4.40	1.35	1.39
1	A	111	TOX	CH2-CZ2	3.46	1.44	1.36
1	B	111	TOX	CH2-CZ3	2.57	1.44	1.38
1	A	111	TOX	CE3-CD2	-2.40	1.37	1.42
1	A	111	TOX	CH2-CZ3	2.31	1.44	1.38
1	A	111	TOX	CZ2-CE2	-2.22	1.36	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	111	TOX	CZ2-CE2-CD2	-4.15	115.29	120.94
1	B	111	TOX	CE3-CD2-CG	-3.04	128.83	134.42
1	A	111	TOX	CZ3-CH2-CZ2	-2.85	116.44	120.44
1	B	111	TOX	CZ3-CE3-CD2	-2.09	117.99	120.89
1	B	111	TOX	CH2-CZ2-CE2	2.05	123.37	119.44

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	A	111	TOX	C-CA-CB-CG
1	B	111	TOX	N-CA-CB-CG
1	B	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 15 ligands modelled in this entry, 2 are monoatomic - leaving 13 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	MPD	B	805	-	7,7,7	0.99	1 (14%)	9,10,10	2.40	3 (33%)
4	OXY	B	804	-	1,1,1	0.08	0	-		
4	OXY	B	803	-	1,1,1	0.67	0	-		
2	HEM	B	801	1	41,50,50	1.88	7 (17%)	45,82,82	1.88	13 (28%)
5	MPD	A	805	-	7,7,7	0.72	0	9,10,10	2.12	4 (44%)
6	TRS	A	807	-	7,7,7	0.97	0	9,9,9	1.13	0
6	TRS	B	806	-	7,7,7	0.71	0	9,9,9	0.98	0
7	PO4	A	808	-	4,4,4	0.66	0	6,6,6	0.93	0
4	OXY	A	803	-	1,1,1	0.12	0	-		
5	MPD	A	806	-	7,7,7	0.67	0	9,10,10	0.86	0
4	OXY	A	804	-	1,1,1	0.50	0	-		
7	PO4	B	807	-	4,4,4	0.71	0	6,6,6	1.79	3 (50%)
2	HEM	A	801	1	41,50,50	1.55	4 (9%)	45,82,82	1.90	14 (31%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	MPD	B	805	-	-	1/5/5/5	-
2	HEM	B	801	1	-	2/12/54/54	-
5	MPD	A	805	-	-	0/5/5/5	-
6	TRS	A	807	-	-	1/9/9/9	-
6	TRS	B	806	-	-	0/9/9/9	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	MPD	A	806	-	-	0/5/5/5	-
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	B	801	HEM	C1B-NB	-7.76	1.26	1.40
2	A	801	HEM	C1B-NB	-4.42	1.32	1.40
2	A	801	HEM	C4D-C3D	3.77	1.51	1.45
2	A	801	HEM	C4D-ND	-3.58	1.34	1.40
2	B	801	HEM	C3C-C2C	-3.30	1.35	1.40
2	B	801	HEM	C4B-NB	-2.81	1.33	1.38
2	B	801	HEM	C1D-ND	-2.67	1.33	1.38
2	B	801	HEM	C2A-C3A	-2.51	1.30	1.37
2	A	801	HEM	O1A-CGA	2.39	1.30	1.22
5	B	805	MPD	C3-C2	-2.35	1.47	1.53
2	B	801	HEM	C3B-C4B	2.24	1.49	1.44
2	B	801	HEM	C1D-C2D	2.01	1.48	1.44

All (37) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	805	MPD	O4-C4-C3	-5.04	91.03	111.36
2	B	801	HEM	CHC-C4B-NB	5.01	129.88	124.43
2	A	801	HEM	C1B-NB-C4B	4.51	109.73	105.07
2	B	801	HEM	C4B-C3B-C2B	-4.40	103.62	107.11
2	A	801	HEM	CHC-C4B-NB	4.34	129.15	124.43
5	A	805	MPD	O4-C4-C3	-3.53	97.11	111.36
2	B	801	HEM	O2A-CGA-CBA	3.53	125.37	114.03
5	B	805	MPD	O2-C2-C3	-3.48	96.73	109.80
2	A	801	HEM	CBA-CAA-C2A	3.30	118.24	112.62
2	B	801	HEM	C1B-NB-C4B	3.22	108.40	105.07
2	A	801	HEM	C4C-CHD-C1D	3.11	126.67	122.56
2	B	801	HEM	CHA-C4D-C3D	-3.07	119.56	125.33
2	A	801	HEM	CMC-C2C-C3C	3.04	130.36	124.68
2	B	801	HEM	CAA-CBA-CGA	-3.04	105.24	113.76
5	A	805	MPD	O2-C2-CM	-2.89	98.80	108.08
2	A	801	HEM	O1D-CGD-CBD	-2.83	113.98	123.08
5	A	805	MPD	CM-C2-C3	2.81	123.03	109.96
2	A	801	HEM	CBD-CAD-C3D	2.75	120.28	112.63
2	A	801	HEM	O2A-CGA-CBA	2.66	122.58	114.03
2	B	801	HEM	CHA-C4D-ND	2.51	127.49	124.38

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	801	HEM	CHD-C1D-ND	2.48	127.13	124.43
2	A	801	HEM	C4B-C3B-C2B	-2.48	105.15	107.11
2	A	801	HEM	CAD-C3D-C4D	2.47	128.98	124.66
2	B	801	HEM	CHC-C4B-C3B	-2.47	120.78	124.57
2	A	801	HEM	CAA-CBA-CGA	-2.45	106.89	113.76
2	B	801	HEM	O2A-CGA-O1A	-2.39	117.35	123.30
5	A	805	MPD	O4-C4-C5	2.37	119.63	109.38
7	B	807	PO4	O4-P-O2	2.34	115.50	107.97
5	B	805	MPD	O2-C2-C1	2.28	115.39	108.08
2	A	801	HEM	C3D-C4D-ND	2.26	112.69	110.17
2	A	801	HEM	O1A-CGA-CBA	-2.22	115.95	123.08
7	B	807	PO4	O3-P-O1	2.16	118.80	110.89
7	B	807	PO4	O4-P-O3	-2.14	101.11	107.97
2	B	801	HEM	CMD-C2D-C1D	2.13	128.28	125.04
2	A	801	HEM	C4D-C3D-C2D	-2.10	103.84	106.90
2	B	801	HEM	CBB-CAB-C3B	-2.04	117.49	127.62
2	B	801	HEM	C4D-C3D-C2D	-2.03	103.94	106.90

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	805	MPD	C2-C3-C4-O4
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
6	A	807	TRS	C1-C-C2-O2

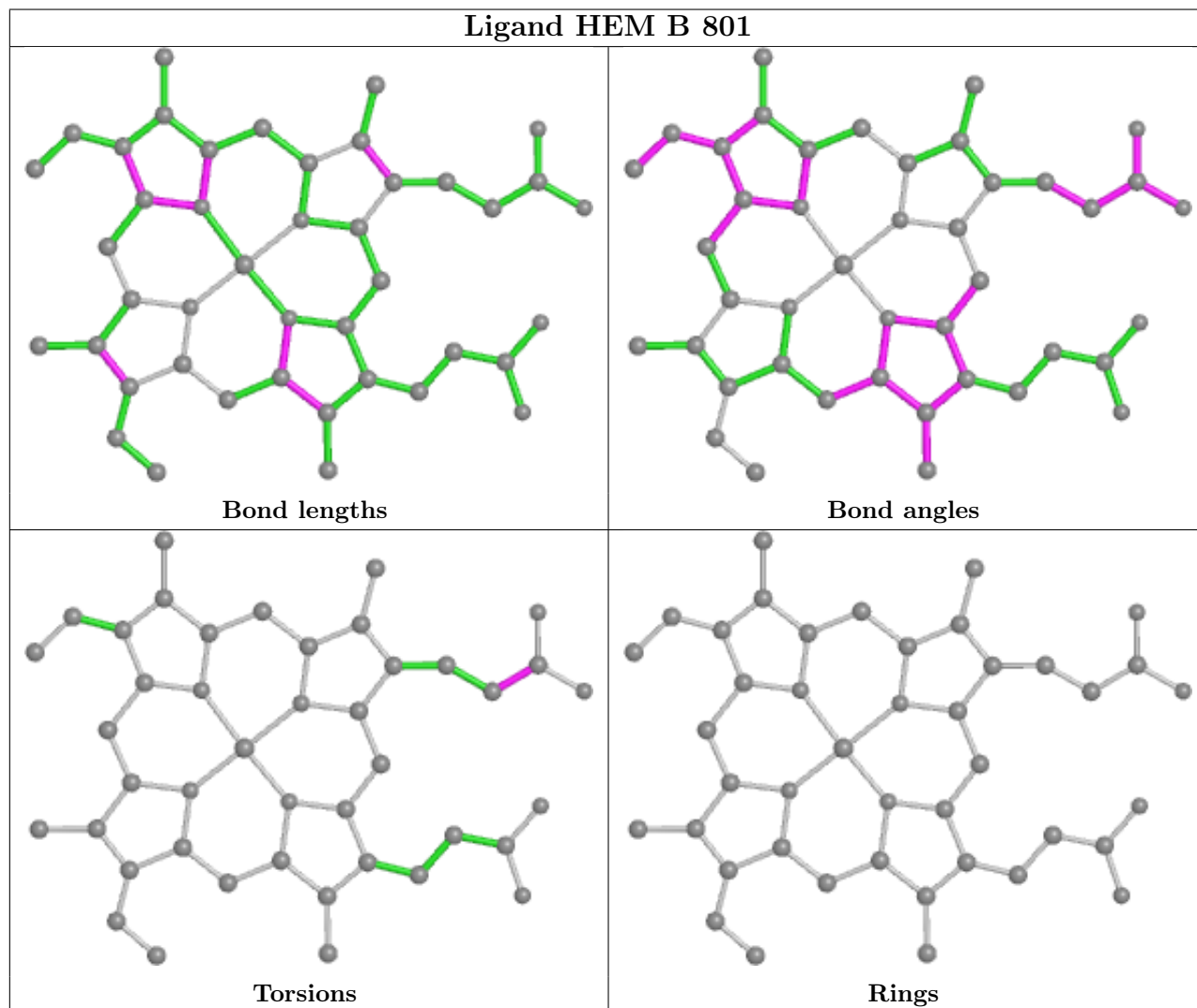
There are no ring outliers.

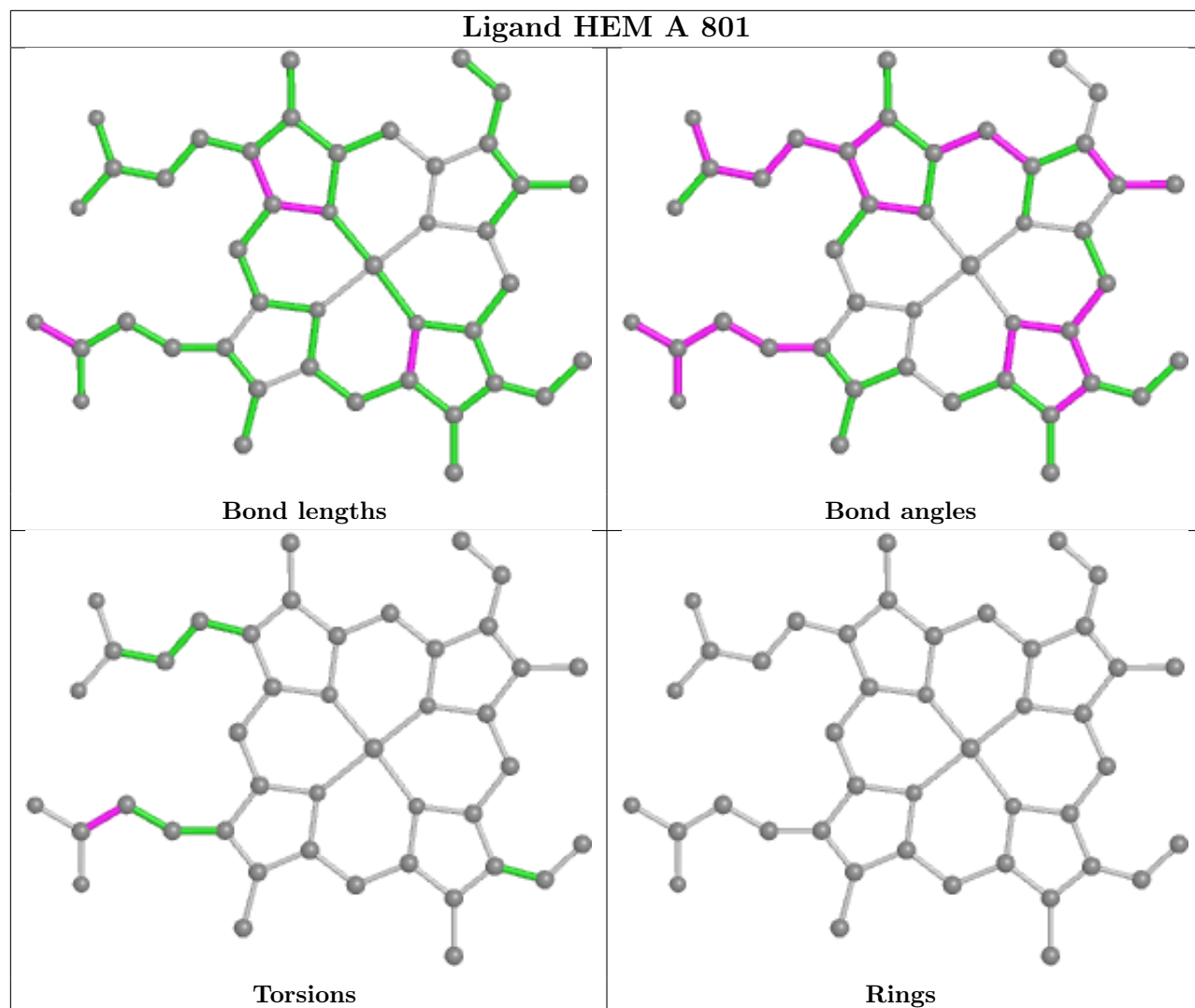
5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	805	MPD	2	0
4	B	803	OXY	1	0
5	A	805	MPD	2	0
5	A	806	MPD	1	0
4	A	804	OXY	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 than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	712/728 (97%)	-0.46	18 (2%) 57 65	15, 23, 47, 81	0
1	B	712/728 (97%)	-0.58	10 (1%) 75 80	14, 21, 40, 79	1 (0%)
All	All	1424/1456 (97%)	-0.52	28 (1%) 65 72	14, 22, 43, 81	1 (0%)

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	748	ALA	4.6
1	B	679	ALA	4.5
1	A	541	GLY	4.4
1	A	222	ASP	3.9
1	B	540	ARG	3.6
1	A	65	ASP	3.3
1	B	748	ALA	3.0
1	A	221	GLY	3.0
1	A	215	PRO	3.0
1	B	65	ASP	2.9
1	A	220	SER	2.9
1	A	679	ALA	2.9
1	A	680	ALA	2.9
1	B	680	ALA	2.9
1	A	224	GLN	2.7
1	A	213	GLY	2.7
1	B	64	LYS	2.7
1	A	540	ARG	2.7
1	A	64	LYS	2.6
1	A	610	ARG	2.6
1	A	608	LYS	2.5
1	B	45	ASN	2.5
1	A	214	GLY	2.4
1	B	44	PRO	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	608	LYS	2.3
1	B	610	ARG	2.3
1	A	378	SER	2.2
1	A	454	ASP	2.2

## 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<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
1	TOX	A	111	16/17	0.96	0.15	17,20,38,39	0
1	TOX	B	111	16/17	0.96	0.14	15,18,32,40	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<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	OXY	A	804	2/2	0.67	0.40	50,50,50,54	0
4	OXY	A	803	2/2	0.82	0.38	59,59,59,61	0
5	MPD	A	806	8/8	0.82	0.22	49,56,62,62	0
5	MPD	A	805	8/8	0.87	0.19	61,65,72,76	0
5	MPD	B	805	8/8	0.89	0.14	53,59,69,74	0
4	OXY	B	803	2/2	0.91	0.12	34,34,34,43	0
4	OXY	B	804	2/2	0.92	0.16	39,39,39,45	0
6	TRS	A	807	8/8	0.93	0.11	29,34,38,38	0
7	PO4	A	808	5/5	0.94	0.28	53,69,70,72	0
7	PO4	B	807	5/5	0.94	0.29	47,58,62,69	0
6	TRS	B	806	8/8	0.96	0.08	23,24,26,26	0
3	NA	B	802	1/1	0.97	0.05	17,17,17,17	0

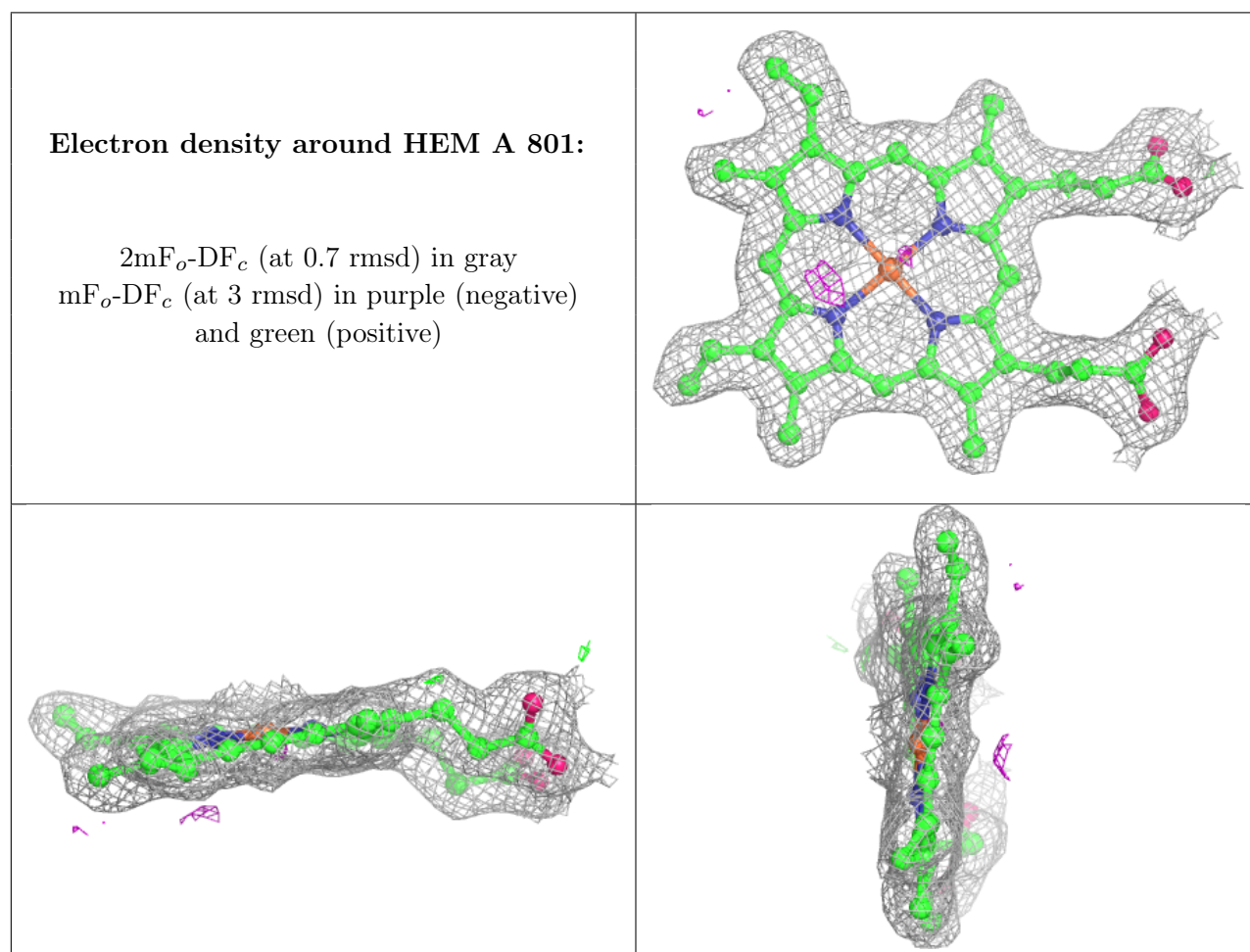
*Continued on next page...*

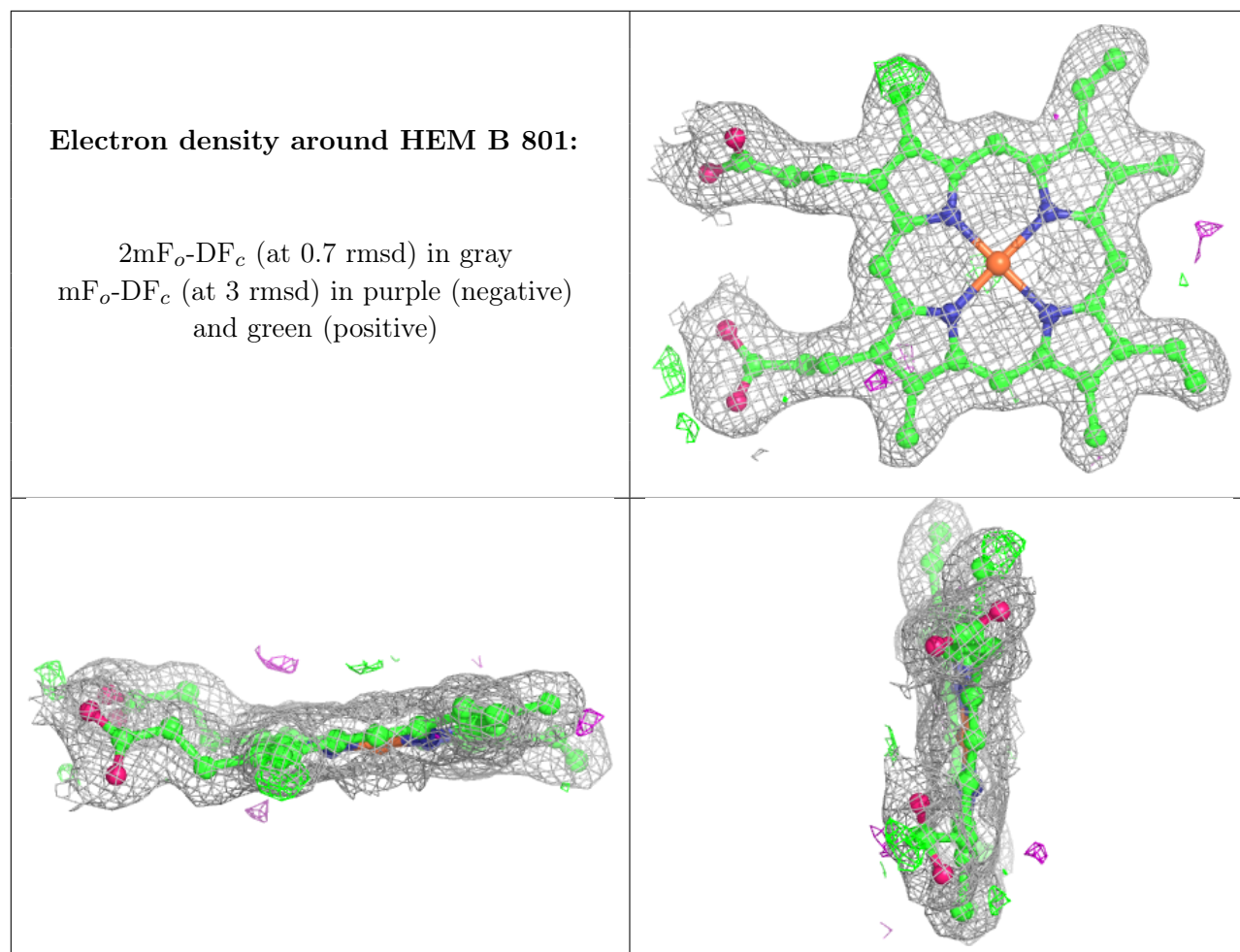


*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	HEM	A	801	43/43	0.98	0.14	16,18,20,22	0
3	NA	A	802	1/1	0.98	0.12	20,20,20,20	0
2	HEM	B	801	43/43	0.99	0.12	15,16,17,20	0

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





## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.