



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 16, 2016 – 01:50 PM EDT

PDB ID : 5SW4  
Title : Crystal structure of native catalase-peroxidase KatG at pH8.0  
Authors : Loewen, P.C.  
Deposited on : 2016-08-08  
Resolution : 1.90 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027939  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027939

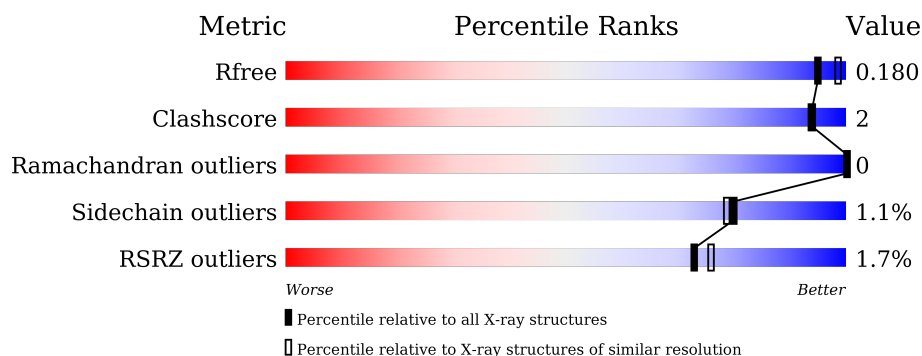
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
$R_{free}$	91344	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	<div> <div>2%</div> <div>88%</div> <div>9%</div> <div>..</div> </div>
1	B	728	<div> <div>2%</div> <div>86%</div> <div>12%</div> <div>.</div> </div>

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
5	PO4	A	804	-	-	-	X
5	PO4	B	803	-	-	-	X
6	MPD	A	805	-	-	-	X
6	MPD	B	804	-	-	-	X
7	TRS	A	807	-	-	-	X
7	TRS	B	805	-	-	-	X

## 2 Entry composition [i](#)

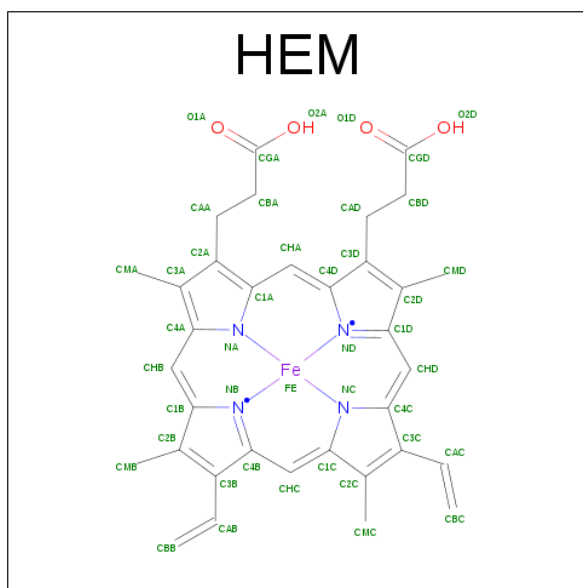
There are 8 unique types of molecules in this entry. The entry contains 12811 atoms, of which 0 are hydrogens and 0 are deuteriums.

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	8	0
			5570	3513	992	1051	14			
1	B	713	Total	C	N	O	S	0	5	0
			5548	3502	989	1043	14			

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

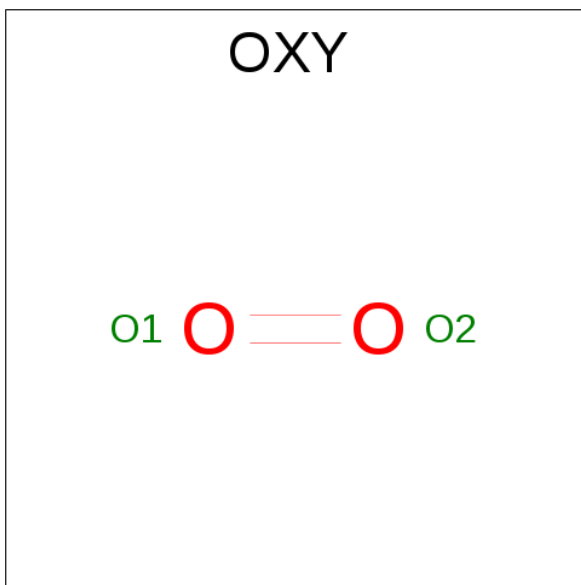


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
2	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

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

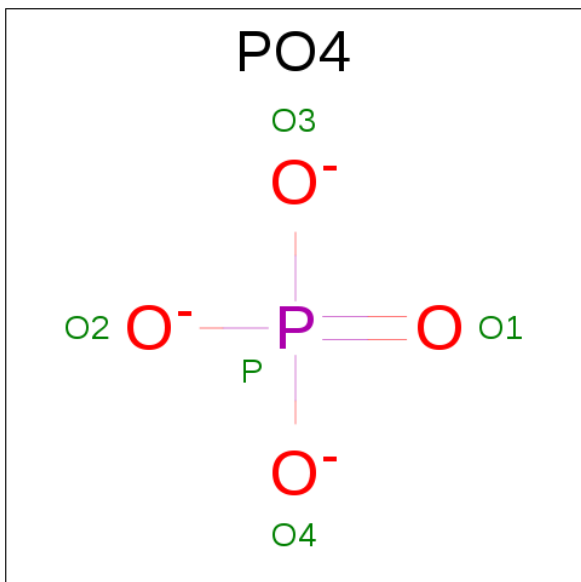
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Na	0	0
			1	1		
3	A	1	Total	Na	0	0
			1	1		

- Molecule 4 is OXYGEN MOLECULE (three-letter code: OXY) (formula: O<sub>2</sub>).



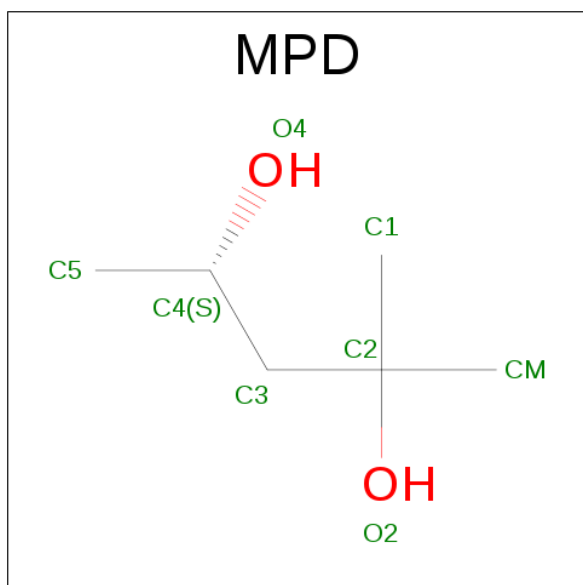
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	O	0	0
			2	2		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



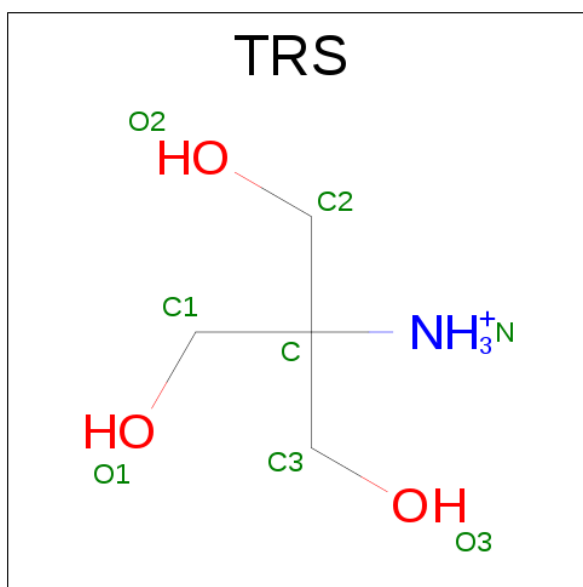
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	P	0	0
			5	4	1		
5	B	1	Total	O	P	0	0
			5	4	1		

- Molecule 6 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula:  $C_6H_{14}O_2$ ).



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

- Molecule 7 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
7	A	1	Total	C	N	O	0	0
			8	4	1	3		
7	B	1	Total	C	N	O	0	0
			8	4	1	3		

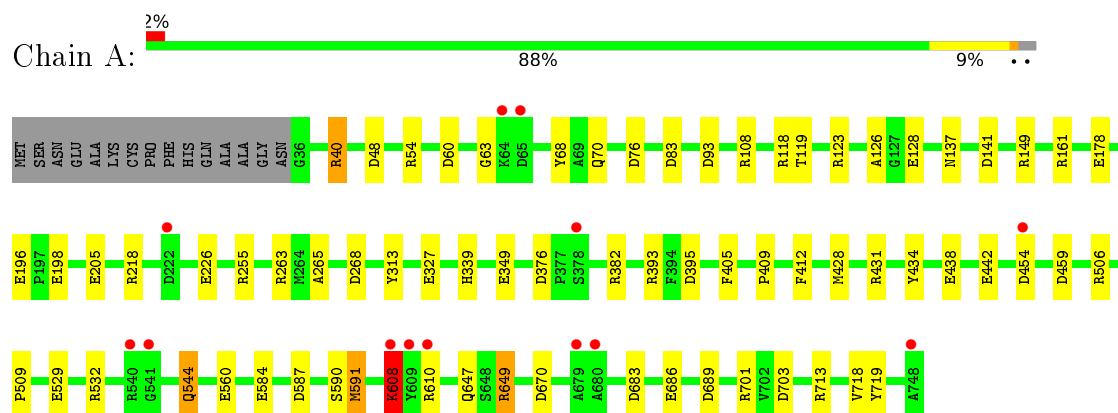
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	762	Total	O	0	0
			762	762		
8	B	791	Total	O	0	0
			791	791		

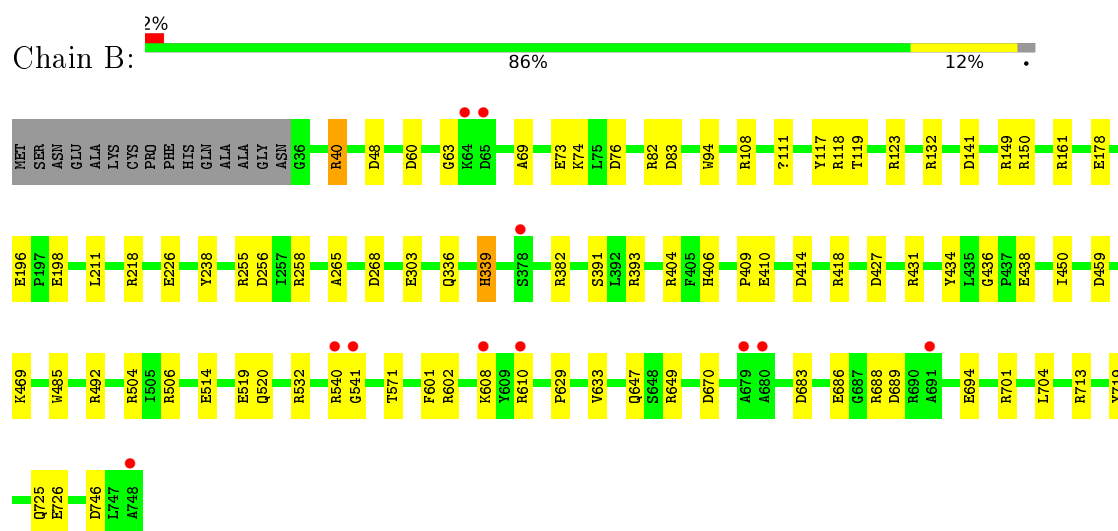
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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, $\alpha$ , $\beta$ , $\gamma$	100.78Å 115.98Å 174.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.90 19.86 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-1.90) 99.9 (19.86-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.84 (at 1.90Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.139 , 0.171 0.153 , 0.180	Depositor DCC
$R_{free}$ test set	15966 reflections (11.01%)	DCC
Wilson B-factor (Å <sup>2</sup> )	20.1	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 53.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	12811	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.99% 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 ⓘ

### 5.1 Standard geometry ⓘ

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

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	1.41	26/5693 (0.5%)	1.33	59/7737 (0.8%)
1	B	1.43	31/5675 (0.5%)	1.24	51/7713 (0.7%)
All	All	1.42	57/11368 (0.5%)	1.29	110/15450 (0.7%)

All (57) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	382	ARG	CZ-NH1	9.86	1.45	1.33
1	A	349	GLU	CD-OE2	-9.31	1.15	1.25
1	B	196	GLU	CD-OE1	-8.61	1.16	1.25
1	B	382	ARG	CZ-NH1	8.56	1.44	1.33
1	B	410	GLU	CG-CD	8.52	1.64	1.51
1	B	519	GLU	CB-CG	8.00	1.67	1.52
1	A	196	GLU	CB-CG	-7.79	1.37	1.52
1	A	178	GLU	CD-OE2	-7.76	1.17	1.25
1	A	198	GLU	CD-OE1	7.63	1.34	1.25
1	B	726	GLU	CG-CD	7.57	1.63	1.51
1	B	436	GLY	C-O	-7.48	1.11	1.23
1	B	601	PHE	CD2-CE2	7.29	1.53	1.39
1	B	694	GLU	CD-OE2	6.83	1.33	1.25
1	B	196	GLU	CD-OE2	-6.80	1.18	1.25
1	A	590	SER	CB-OG	-6.74	1.33	1.42
1	B	196	GLU	CB-CG	-6.54	1.39	1.52
1	A	544	GLN	CD-OE1	6.50	1.38	1.24
1	A	327	GLU	CD-OE1	6.38	1.32	1.25
1	B	391	SER	CA-CB	6.22	1.62	1.52
1	A	438	GLU	CD-OE1	-6.08	1.19	1.25
1	B	63	GLY	N-CA	6.01	1.55	1.46
1	A	205	GLU	CD-OE2	5.98	1.32	1.25
1	B	198	GLU	CD-OE1	5.92	1.32	1.25
1	A	126	ALA	C-O	5.89	1.34	1.23

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	68	TYR	CE1-CZ	5.89	1.46	1.38
1	A	313	TYR	CE2-CZ	5.86	1.46	1.38
1	A	226	GLU	CG-CD	5.81	1.60	1.51
1	B	434	TYR	CG-CD2	-5.78	1.31	1.39
1	B	226	GLU	CG-CD	5.77	1.60	1.51
1	B	336	GLN	CG-CD	5.77	1.64	1.51
1	B	514	GLU	CD-OE2	-5.75	1.19	1.25
1	A	703	ASP	CG-OD2	5.72	1.38	1.25
1	B	746	ASP	CB-CG	5.62	1.63	1.51
1	B	117	TYR	CE2-CZ	5.61	1.45	1.38
1	A	128	GLU	CG-CD	5.55	1.60	1.51
1	A	584	GLU	CG-CD	5.54	1.60	1.51
1	B	178	GLU	CD-OE2	5.47	1.31	1.25
1	A	560	GLU	CD-OE2	-5.42	1.19	1.25
1	A	428	MET	CG-SD	-5.41	1.67	1.81
1	B	519	GLU	CG-CD	-5.41	1.43	1.51
1	B	686	GLU	CG-CD	5.37	1.59	1.51
1	B	438	GLU	CD-OE1	-5.36	1.19	1.25
1	B	725	GLN	CD-NE2	5.35	1.46	1.32
1	B	485	TRP	CE3-CZ3	5.33	1.47	1.38
1	A	442	GLU	CD-OE2	-5.33	1.19	1.25
1	B	303	GLU	CD-OE2	-5.31	1.19	1.25
1	A	63	GLY	N-CA	5.31	1.54	1.46
1	B	94	TRP	CE3-CZ3	5.24	1.47	1.38
1	A	719	TYR	CE1-CZ	5.20	1.45	1.38
1	B	196	GLU	CG-CD	5.18	1.59	1.51
1	B	434	TYR	CE2-CZ	-5.18	1.31	1.38
1	A	686	GLU	CG-CD	5.15	1.59	1.51
1	A	137	ASN	N-CA	-5.11	1.36	1.46
1	A	196	GLU	CG-CD	5.05	1.59	1.51
1	B	161	ARG	CZ-NH1	5.04	1.39	1.33
1	B	602	ARG	CZ-NH2	-5.02	1.26	1.33
1	A	532	ARG	CD-NE	-5.02	1.38	1.46

All (110) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	255[A]	ARG	NE-CZ-NH2	-21.90	109.35	120.30
1	A	255[B]	ARG	NE-CZ-NH2	-21.90	109.35	120.30
1	A	255[A]	ARG	NE-CZ-NH1	21.63	131.12	120.30
1	A	255[B]	ARG	NE-CZ-NH1	21.63	131.12	120.30
1	B	82	ARG	NE-CZ-NH2	-11.45	114.58	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	591	MET	CG-SD-CE	10.89	117.63	100.20
1	A	382	ARG	NE-CZ-NH2	-10.02	115.29	120.30
1	B	76	ASP	CB-CG-OD2	-10.01	109.29	118.30
1	B	704	LEU	CB-CG-CD2	9.93	127.89	111.00
1	A	54	ARG	NE-CZ-NH1	9.85	125.22	120.30
1	A	713	ARG	NE-CZ-NH1	9.54	125.07	120.30
1	B	393	ARG	NE-CZ-NH2	-9.45	115.57	120.30
1	A	218	ARG	NE-CZ-NH2	-9.02	115.79	120.30
1	B	382	ARG	NE-CZ-NH2	-9.01	115.80	120.30
1	A	123	ARG	NE-CZ-NH1	8.94	124.77	120.30
1	B	108	ARG	NE-CZ-NH2	-8.93	115.84	120.30
1	A	649	ARG	NE-CZ-NH1	8.77	124.68	120.30
1	A	76	ASP	CB-CG-OD2	-8.58	110.58	118.30
1	B	701	ARG	NE-CZ-NH1	8.51	124.56	120.30
1	A	40	ARG	NE-CZ-NH1	8.41	124.51	120.30
1	B	683	ASP	CB-CG-OD1	8.30	125.77	118.30
1	A	434	TYR	CB-CG-CD1	8.29	125.98	121.00
1	A	108	ARG	NE-CZ-NH1	8.24	124.42	120.30
1	A	683	ASP	CB-CG-OD1	8.22	125.70	118.30
1	B	649	ARG	NE-CZ-NH1	8.03	124.32	120.30
1	A	268	ASP	CB-CG-OD2	-8.01	111.09	118.30
1	B	683	ASP	CB-CG-OD2	-7.88	111.21	118.30
1	A	649	ARG	NE-CZ-NH2	-7.84	116.38	120.30
1	B	713	ARG	NE-CZ-NH1	7.76	124.18	120.30
1	A	255[A]	ARG	CD-NE-CZ	7.76	134.46	123.60
1	A	255[B]	ARG	CD-NE-CZ	7.76	134.46	123.60
1	B	268	ASP	CB-CG-OD1	7.74	125.26	118.30
1	B	268	ASP	CB-CG-OD2	-7.67	111.39	118.30
1	B	82	ARG	NE-CZ-NH1	7.61	124.11	120.30
1	B	404	ARG	NE-CZ-NH2	-7.61	116.49	120.30
1	A	54	ARG	NE-CZ-NH2	-7.59	116.50	120.30
1	A	670	ASP	CB-CG-OD2	-7.47	111.58	118.30
1	A	382	ARG	NE-CZ-NH1	7.38	123.99	120.30
1	B	532	ARG	NE-CZ-NH2	-7.34	116.63	120.30
1	B	492	ARG	NE-CZ-NH1	7.24	123.92	120.30
1	B	382	ARG	NE-CZ-NH1	7.08	123.84	120.30
1	B	40	ARG	NE-CZ-NH1	7.05	123.83	120.30
1	B	689	ASP	CB-CG-OD1	7.04	124.64	118.30
1	B	123	ARG	NE-CZ-NH1	6.98	123.79	120.30
1	A	701	ARG	NE-CZ-NH1	6.98	123.79	120.30
1	A	60	ASP	CB-CG-OD1	6.90	124.51	118.30
1	B	726	GLU	OE1-CD-OE2	-6.90	115.02	123.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	434	TYR	CB-CG-CD1	6.86	125.11	121.00
1	A	149	ARG	NE-CZ-NH1	6.83	123.71	120.30
1	A	532	ARG	NE-CZ-NH2	-6.74	116.93	120.30
1	B	258	ARG	NE-CZ-NH1	6.72	123.66	120.30
1	B	506	ARG	NE-CZ-NH1	6.72	123.66	120.30
1	A	587	ASP	CB-CG-OD2	-6.72	112.25	118.30
1	B	610	ARG	NE-CZ-NH1	6.63	123.61	120.30
1	A	529	GLU	OE1-CD-OE2	-6.51	115.49	123.30
1	B	108	ARG	NE-CZ-NH1	6.49	123.55	120.30
1	A	48	ASP	CB-CG-OD1	6.48	124.13	118.30
1	A	689	ASP	CB-CG-OD2	-6.45	112.50	118.30
1	A	54	ARG	CG-CD-NE	-6.44	98.27	111.80
1	A	93	ASP	CB-CG-OD1	6.41	124.07	118.30
1	A	83	ASP	CB-CG-OD2	6.39	124.05	118.30
1	A	713	ARG	NE-CZ-NH2	-6.35	117.12	120.30
1	A	506	ARG	NE-CZ-NH1	6.34	123.47	120.30
1	B	83	ASP	CB-CG-OD2	6.32	123.98	118.30
1	B	504	ARG	NE-CZ-NH2	-6.28	117.16	120.30
1	A	60	ASP	CB-CG-OD2	-6.20	112.72	118.30
1	B	649	ARG	NE-CZ-NH2	-6.13	117.24	120.30
1	B	256	ASP	CB-CG-OD1	6.11	123.80	118.30
1	B	149	ARG	NE-CZ-NH1	6.09	123.34	120.30
1	A	405	PHE	CB-CG-CD1	-6.07	116.55	120.80
1	A	584	GLU	OE1-CD-OE2	-6.02	116.08	123.30
1	B	218	ARG	NE-CZ-NH2	-6.00	117.30	120.30
1	B	670	ASP	CB-CG-OD1	5.97	123.67	118.30
1	A	263	ARG	NE-CZ-NH2	-5.92	117.34	120.30
1	A	670	ASP	CB-CG-OD1	5.90	123.61	118.30
1	A	689	ASP	CB-CG-OD1	5.89	123.60	118.30
1	B	541	GLY	N-CA-C	-5.89	98.38	113.10
1	A	70	GLN	CA-CB-CG	5.88	126.34	113.40
1	A	587	ASP	CB-CG-OD1	5.88	123.59	118.30
1	B	504	ARG	NE-CZ-NH1	5.81	123.21	120.30
1	B	450	ILE	CA-CB-CG1	-5.79	100.00	111.00
1	A	412	PHE	CB-CG-CD2	-5.75	116.77	120.80
1	B	459	ASP	CB-CG-OD2	-5.72	113.15	118.30
1	A	608	LYS	CB-CG-CD	5.68	126.38	111.60
1	A	459	ASP	CB-CG-OD2	-5.65	113.21	118.30
1	A	610	ARG	NE-CZ-NH1	5.65	123.12	120.30
1	B	48	ASP	CB-CG-OD1	5.65	123.38	118.30
1	A	393	ARG	NE-CZ-NH2	-5.63	117.48	120.30
1	A	376	ASP	CB-CG-OD2	-5.61	113.26	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	128	GLU	OE1-CD-OE2	-5.60	116.58	123.30
1	A	395	ASP	CB-CG-OD1	5.59	123.33	118.30
1	A	718	VAL	CG1-CB-CG2	-5.57	102.00	110.90
1	B	132	ARG	NE-CZ-NH2	-5.56	117.52	120.30
1	B	688	ARG	NE-CZ-NH2	-5.50	117.55	120.30
1	A	532	ARG	NE-CZ-NH1	5.48	123.04	120.30
1	B	519	GLU	CG-CD-OE1	-5.46	107.38	118.30
1	A	431	ARG	NE-CZ-NH1	-5.42	117.59	120.30
1	B	60	ASP	CB-CG-OD1	5.34	123.11	118.30
1	A	218	ARG	NE-CZ-NH1	5.34	122.97	120.30
1	B	689	ASP	CB-CG-OD2	-5.32	113.51	118.30
1	B	150	ARG	NE-CZ-NH2	-5.30	117.65	120.30
1	B	414	ASP	CB-CG-OD2	5.29	123.06	118.30
1	A	454	ASP	CB-CG-OD2	5.28	123.05	118.30
1	B	161	ARG	CG-CD-NE	-5.27	100.73	111.80
1	A	268	ASP	CB-CG-OD1	5.24	123.01	118.30
1	B	418	ARG	NE-CZ-NH1	5.18	122.89	120.30
1	A	68	TYR	CG-CD1-CE1	5.16	125.42	121.30
1	B	427	ASP	CB-CG-OD1	-5.16	113.66	118.30
1	B	339[A]	HIS	CB-CA-C	5.08	120.57	110.40
1	B	339[B]	HIS	CB-CA-C	5.08	120.57	110.40

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5570	0	5356	10	0
1	B	5548	0	5354	22	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	1	0
5	A	5	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	5	0	0	0	0
6	A	16	0	28	5	0
6	B	8	0	14	2	0
7	A	8	0	12	0	0
7	B	8	0	11	0	0
8	A	762	0	0	7	0
8	B	791	0	0	10	0
All	All	12811	0	10835	40	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 (40) 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:B:339[B]:HIS:HE1	1:B:406:HIS:HA	1.38	0.89
1:B:520:GLN:HG3	8:B:1383:HOH:O	1.73	0.87
1:B:647:GLN:HG2	8:B:1050:HOH:O	1.75	0.86
1:B:571:THR:HG23	8:B:1239:HOH:O	1.82	0.80
1:A:339[B]:HIS:CE1	8:A:925:HOH:O	2.33	0.80
1:B:431:ARG:CD	8:B:1510:HOH:O	2.33	0.76
1:B:69:ALA:O	1:B:73:GLU:HG2	1.86	0.74
1:B:339[B]:HIS:CE1	1:B:406:HIS:HA	2.23	0.72
1:A:339[B]:HIS:CD2	1:A:409:PRO:HB3	2.26	0.71
1:B:431:ARG:HD2	8:B:1510:HOH:O	1.90	0.71
1:B:255[A]:ARG:HD3	8:B:1370:HOH:O	1.91	0.70
6:A:805:MPD:H51	8:A:1524:HOH:O	1.98	0.64
1:B:255[B]:ARG:HG2	8:B:1633:HOH:O	1.99	0.62
6:A:805:MPD:O4	6:A:805:MPD:CM	2.47	0.62
1:B:540:ARG:NE	1:B:540:ARG:HA	2.17	0.59
6:B:804:MPD:O4	6:B:804:MPD:H12	2.03	0.59
1:B:119[A]:THR:CG2	1:B:265:ALA:HB2	2.32	0.59
1:B:119[A]:THR:HG22	1:B:265:ALA:HB2	1.86	0.57
1:A:119[A]:THR:CG2	1:A:265:ALA:HB2	2.35	0.56
1:A:119[A]:THR:HG22	1:A:265:ALA:HB2	1.90	0.54
1:B:633[B]:VAL:CG1	1:B:719:TYR:CZ	2.91	0.54
1:B:339[A]:HIS:CD2	1:B:409:PRO:HB3	2.47	0.50
6:A:805:MPD:HM1	6:A:805:MPD:O4	2.12	0.50
1:A:339[B]:HIS:HE1	8:A:925:HOH:O	1.82	0.49
1:B:111:TOX:H9	1:B:238:TYR:OH	2.13	0.49
4:A:803:OXY:O2	8:A:901:HOH:O	2.20	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:629:PRO:O	1:B:633[A]:VAL:HG23	2.14	0.47
6:A:806:MPD:HM2	8:A:1476:HOH:O	2.16	0.45
1:A:608:LYS:HE2	1:A:608:LYS:HA	1.98	0.45
6:B:804:MPD:O4	6:B:804:MPD:C1	2.60	0.44
1:B:255[B]:ARG:CG	8:B:1633:HOH:O	2.64	0.43
1:A:509:PRO:HD2	1:A:591:MET:HG2	2.00	0.43
6:A:806:MPD:CM	8:A:1476:HOH:O	2.67	0.43
1:B:211:LEU:HB2	8:B:1469:HOH:O	2.18	0.43
1:B:469:LYS:HD3	8:B:1385:HOH:O	2.19	0.41
1:A:339[A]:HIS:HB2	8:A:1442:HOH:O	2.20	0.41
1:A:339[B]:HIS:CG	1:A:409:PRO:HB3	2.54	0.41
1:B:339[B]:HIS:HE1	1:B:406:HIS:CA	2.21	0.41
1:A:647:GLN:CD	1:A:647:GLN:N	2.74	0.41
1:B:633[A]:VAL:HG22	1:B:719:TYR:CZ	2.55	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%)	709 (99%)	8 (1%)	0	100	100
1	B	715/728 (98%)	704 (98%)	11 (2%)	0	100	100
All	All	1432/1456 (98%)	1413 (99%)	19 (1%)	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%)	549 (99%)	7 (1%)	76	73
1	B	554/560 (99%)	549 (99%)	5 (1%)	84	83
All	All	1110/1120 (99%)	1098 (99%)	12 (1%)	80	79

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	40	ARG
1	A	118	ARG
1	A	141	ASP
1	A	161	ARG
1	A	544	GLN
1	A	608	LYS
1	A	649	ARG
1	B	40	ARG
1	B	74	LYS
1	B	118	ARG
1	B	141	ASP
1	B	608	LYS

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

Mol	Chain	Res	Type
1	A	70	GLN
1	A	247	ASN
1	B	46	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

3 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[A]	2	10,17,18	2.82	5 (50%)	9,23,25	1.94	4 (44%)
1	TOX	A	111[B]	-	10,17,18	2.82	5 (50%)	9,23,25	1.94	4 (44%)
1	TOX	B	111	1,2	10,17,18	2.26	4 (40%)	9,23,25	2.16	3 (33%)

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[A]	2	-	0/3/8/10	0/2/2/2
1	TOX	A	111[B]	-	-	0/3/8/10	0/2/2/2
1	TOX	B	111	1,2	-	0/3/8/10	0/2/2/2

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	111[A]	TOX	CD1-NE1	-7.10	1.32	1.39
1	A	111[B]	TOX	CD1-NE1	-7.10	1.32	1.39
1	B	111	TOX	CZ2-CE2	-3.87	1.33	1.41
1	B	111	TOX	CD1-NE1	-2.03	1.37	1.39
1	A	111[A]	TOX	CD2-CE2	2.22	1.44	1.41
1	A	111[B]	TOX	CD2-CE2	2.22	1.44	1.41
1	A	111[A]	TOX	CH2-CZ2	2.29	1.41	1.36
1	A	111[B]	TOX	CH2-CZ2	2.29	1.41	1.36
1	A	111[A]	TOX	O-C	2.55	1.31	1.19
1	A	111[B]	TOX	O-C	2.55	1.31	1.19
1	A	111[A]	TOX	CZ3-CE3	2.69	1.42	1.36
1	A	111[B]	TOX	CZ3-CE3	2.69	1.42	1.36
1	B	111	TOX	CD2-CE2	3.57	1.46	1.41
1	B	111	TOX	O-C	3.57	1.36	1.19

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	111	TOX	CB-CG-CD1	-5.06	121.72	127.97
1	A	111[A]	TOX	O-C-CA	-3.66	115.92	125.72

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	111[B]	TOX	O-C-CA	-3.66	115.92	125.72
1	A	111[A]	TOX	CZ2-CE2-CD2	-2.41	117.55	120.58
1	A	111[B]	TOX	CZ2-CE2-CD2	-2.41	117.55	120.58
1	B	111	TOX	O-C-CA	-2.39	119.31	125.72
1	B	111	TOX	CZ2-CE2-CD2	-2.35	117.63	120.58
1	A	111[A]	TOX	CB-CG-CD1	-2.21	125.23	127.97
1	A	111[B]	TOX	CB-CG-CD1	-2.21	125.23	127.97
1	A	111[A]	TOX	CE3-CD2-CE2	2.45	122.30	119.83
1	A	111[B]	TOX	CE3-CD2-CE2	2.45	122.30	119.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 2 are monoatomic - leaving 10 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	HEM	A	801	1	24,50,50	1.38	4 (16%)	16,82,82	1.94	5 (31%)
4	OXY	A	803	-	1,1,1	0.15	0	0,0,0	0.00	-
5	PO4	A	804	-	4,4,4	1.15	1 (25%)	6,6,6	0.34	0
6	MPD	A	805	-	6,7,7	0.90	0	6,10,10	1.64	1 (16%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	MPD	A	806	-	6,7,7	0.89	0	6,10,10	1.19	1 (16%)
7	TRS	A	807	-	7,7,7	2.10	2 (28%)	9,9,9	3.38	5 (55%)
2	HEM	B	801	1	24,50,50	1.53	4 (16%)	16,82,82	2.77	9 (56%)
5	PO4	B	803	-	4,4,4	1.15	1 (25%)	6,6,6	0.40	0
6	MPD	B	804	-	6,7,7	0.71	0	6,10,10	1.46	1 (16%)
7	TRS	B	805	-	7,7,7	1.74	2 (28%)	9,9,9	1.98	1 (11%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEM	A	801	1	-	0/6/54/54	0/0/8/8
4	OXY	A	803	-	-	0/0/0/0	0/0/0/0
5	PO4	A	804	-	-	0/0/0/0	0/0/0/0
6	MPD	A	805	-	-	0/5/5/5	0/0/0/0
6	MPD	A	806	-	-	0/5/5/5	0/0/0/0
7	TRS	A	807	-	-	0/9/9/9	0/0/0/0
2	HEM	B	801	1	-	0/6/54/54	0/0/8/8
5	PO4	B	803	-	-	0/0/0/0	0/0/0/0
6	MPD	B	804	-	-	0/5/5/5	0/0/0/0
7	TRS	B	805	-	-	0/9/9/9	0/0/0/0

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	HEM	C1B-NB	-2.96	1.32	1.36
2	B	801	HEM	C3C-C2C	-2.94	1.36	1.40
2	B	801	HEM	C3B-C2B	-2.92	1.36	1.40
2	A	801	HEM	C4D-ND	-2.60	1.33	1.36
2	A	801	HEM	CAD-C3D	-2.54	1.48	1.52
7	B	805	TRS	O3-C3	-2.31	1.34	1.42
5	B	803	PO4	P-O4	2.02	1.60	1.53
7	A	807	TRS	O1-C1	2.16	1.49	1.42
5	A	804	PO4	P-O1	2.20	1.61	1.53
2	B	801	HEM	CMA-C3A	2.30	1.56	1.51
7	B	805	TRS	C1-C	2.64	1.56	1.53
2	A	801	HEM	C4C-NC	3.10	1.40	1.36
2	B	801	HEM	C4C-NC	3.65	1.41	1.36
7	A	807	TRS	C1-C	4.41	1.59	1.53

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	801	HEM	C3B-CAB-CBB	-5.27	115.81	126.40
2	A	801	HEM	CAA-CBA-CGA	-4.65	103.73	112.78
2	B	801	HEM	CAA-CBA-CGA	-4.16	104.70	112.78
7	A	807	TRS	C3-C-C2	-4.01	102.20	110.65
2	B	801	HEM	CMA-C3A-C4A	-3.25	122.78	128.31
2	A	801	HEM	C3B-CAB-CBB	-2.57	121.24	126.40
6	A	806	MPD	CM-C2-C1	-2.55	104.33	110.41
2	B	801	HEM	CAA-C2A-C3A	-2.13	122.92	129.00
7	A	807	TRS	C2-C-C1	2.01	114.88	110.65
6	B	804	MPD	C1-C2-C3	2.09	121.16	109.98
2	A	801	HEM	CAD-CBD-CGD	2.21	117.09	112.78
2	B	801	HEM	CMD-C2D-C3D	2.23	129.90	125.24
6	A	805	MPD	O4-C4-C5	2.39	121.38	109.47
7	A	807	TRS	O3-C3-C	2.64	117.18	110.92
2	A	801	HEM	CBD-CAD-C3D	2.74	117.28	112.47
2	A	801	HEM	CBA-CAA-C2A	2.83	117.47	112.49
2	B	801	HEM	CMC-C2C-C3C	2.90	130.77	125.09
2	B	801	HEM	CMA-C3A-C2A	3.20	131.94	125.24
2	B	801	HEM	CBD-CAD-C3D	3.37	118.39	112.47
7	A	807	TRS	C3-C-C1	3.96	119.00	110.65
2	B	801	HEM	CAD-CBD-CGD	4.15	120.86	112.78
7	B	805	TRS	O1-C1-C	5.22	123.32	110.92
7	A	807	TRS	O1-C1-C	7.34	128.34	110.92

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	803	OXY	1	0
6	A	805	MPD	3	0
6	A	806	MPD	2	0
6	B	804	MPD	2	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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	13 (1%) 71 74	14, 21, 37, 74	0
1	B	712/728 (97%)	-0.54	11 (1%) 76 79	13, 19, 34, 68	0
All	All	1424/1456 (97%)	-0.50	24 (1%) 73 76	13, 20, 36, 74	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	748	ALA	6.2
1	B	748	ALA	4.5
1	B	679	ALA	4.3
1	A	540	ARG	3.9
1	B	610	ARG	3.6
1	A	610	ARG	3.5
1	A	541	GLY	3.3
1	B	680	ALA	3.3
1	B	65	ASP	3.2
1	B	64	LYS	3.2
1	B	540	ARG	3.2
1	B	608	LYS	3.1
1	A	64	LYS	3.0
1	B	541	GLY	3.0
1	A	608	LYS	3.0
1	A	65	ASP	2.7
1	A	454	ASP	2.6
1	A	222	ASP	2.5
1	A	378	SER	2.3
1	B	378	SER	2.2
1	A	679	ALA	2.2
1	B	691	ALA	2.2
1	A	609	TYR	2.1
1	A	680	ALA	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	TOX	A	111[A]	16/17	0.94	0.10	-	13,17,31,32	0
1	TOX	B	111	16/17	0.97	0.07	-	13,16,26,28	0
1	TOX	A	111[B]	16/17	0.94	0.10	-	13,17,24,32	1

## 6.3 Carbohydrates [i](#)

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	PO4	B	803	5/5	0.88	0.24	6.84	56,63,68,78	0
6	MPD	A	805	8/8	0.84	0.19	5.06	48,55,70,71	0
5	PO4	A	804	5/5	0.77	0.23	4.32	59,67,76,89	0
7	TRS	A	807	8/8	0.91	0.17	4.07	30,37,43,52	0
6	MPD	B	804	8/8	0.91	0.13	3.29	46,53,60,61	0
7	TRS	B	805	8/8	0.95	0.11	2.30	24,31,33,34	0
3	NA	A	802	1/1	1.00	0.08	1.15	17,17,17,17	0
3	NA	B	802	1/1	0.99	0.07	-0.02	16,16,16,16	0
2	HEM	B	801	43/43	0.99	0.06	-0.99	13,15,16,18	0
2	HEM	A	801	43/43	0.99	0.05	-1.30	14,16,18,19	0
4	OXY	A	803	2/2	0.98	0.35	-	31,31,31,37	0
6	MPD	A	806	8/8	0.88	0.17	-	43,45,56,60	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.