W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 15, 2017 – 04:36 am GMT

PDB ID	:	1GGH
Title	:	CRYSTAL STRUCTURE OF CATALASE HPII FROM ESCHERICHIA COLI, HIS128ALA VARIANT.
Authors	:	Melik-Adamyan, W.R.; Bravo, J.; Carpena, X.; Switala, J.; Mate, M.J.; Fita, I.; Loewen, P.C.
Deposited on	:	2000-08-21
Resolution	:	2.15 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

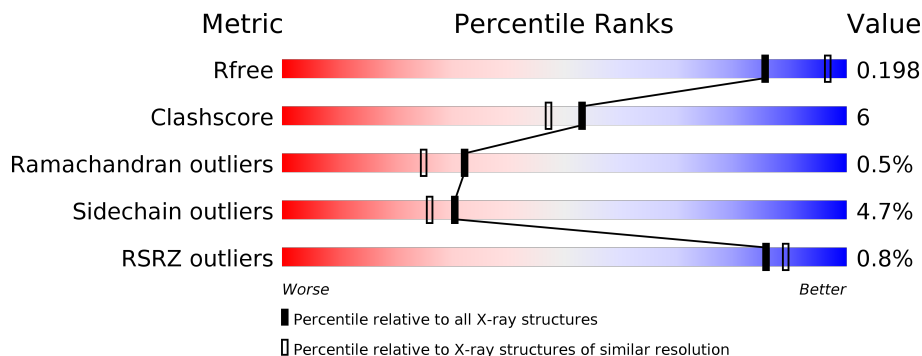|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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	Percentile Ranks	Value
Rfree		0.198
Clashscore		6
Ramachandran outliers		0.5%
Sidechain outliers		4.7%
RSRZ outliers		0.8%

Worse ← → Better

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	1170 (2.16-2.16)
Clashscore	112137	1278 (2.16-2.16)
Ramachandran outliers	110173	1256 (2.16-2.16)
Sidechain outliers	110143	1255 (2.16-2.16)
RSRZ outliers	101464	1175 (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 on the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	753	81% 14% · ·
1	B	753	79% 15% · ·
1	C	753	77% 17% · ·
1	D	753	79% 16% · · ·

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 25922 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 HPII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	727	Total	C	N	O	S	0	0	0
			5741	3644	1003	1082	12			
1	B	727	Total	C	N	O	S	0	0	0
			5741	3644	1003	1082	12			
1	C	727	Total	C	N	O	S	0	0	0
			5741	3644	1003	1082	12			
1	D	727	Total	C	N	O	S	0	0	0
			5741	3644	1003	1082	12			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	128	ALA	HIS	ENGINEERED	UNP P21179
B	128	ALA	HIS	ENGINEERED	UNP P21179
C	128	ALA	HIS	ENGINEERED	UNP P21179
D	128	ALA	HIS	ENGINEERED	UNP P21179

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	737	Total	O	0	0
			737	737		
3	B	661	Total	O	0	0
			661	661		
3	C	675	Total	O	0	0
			675	675		
3	D	713	Total	O	0	0
			713	713		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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.
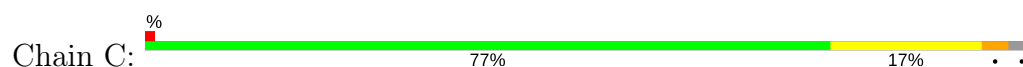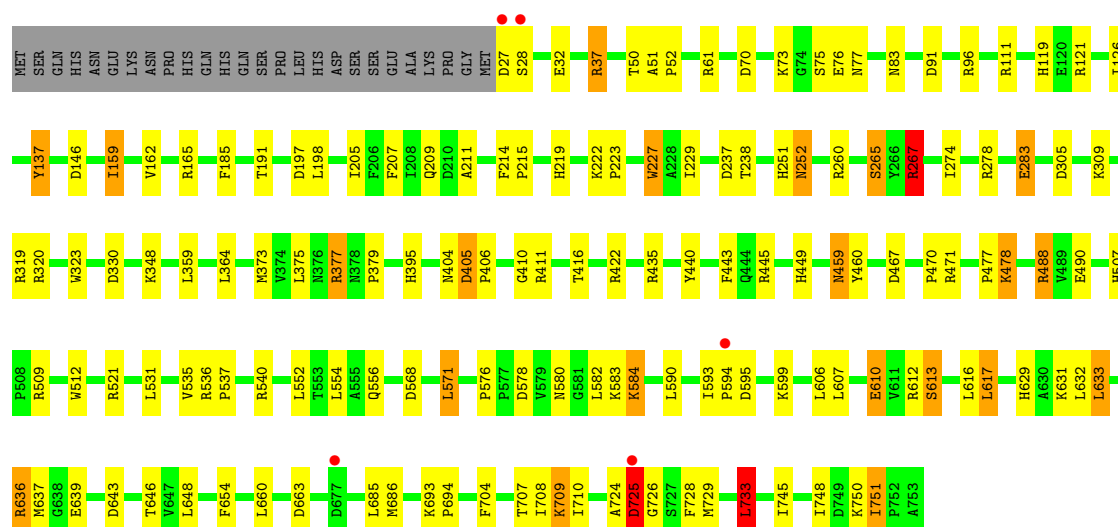
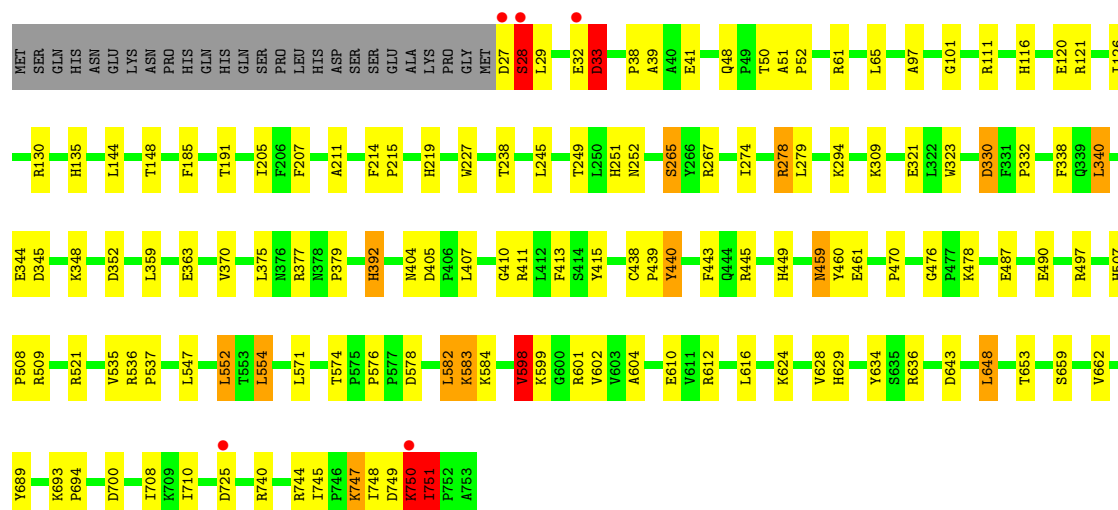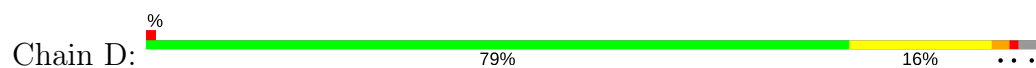● Molecule 1: CATALASE HPII
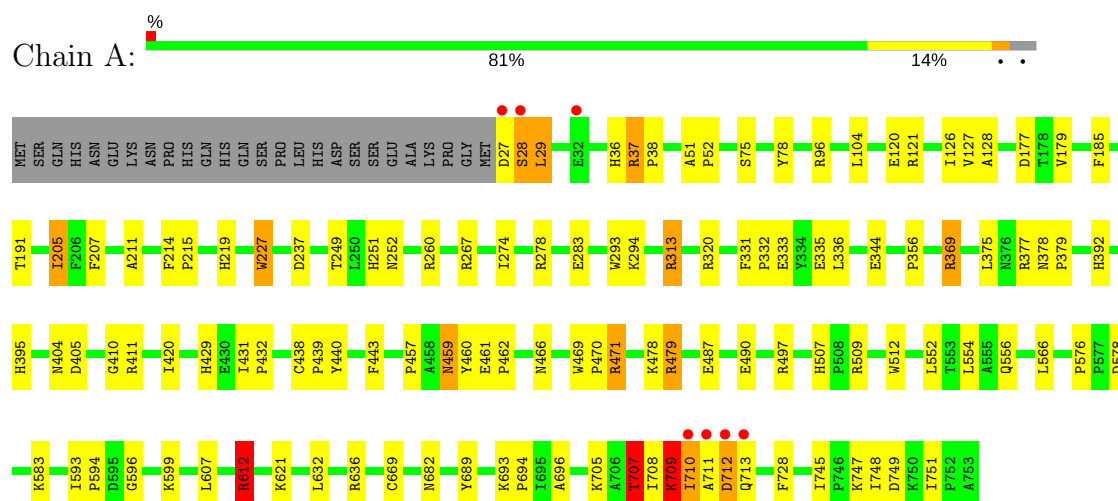
Chain A:

● Molecule 1: CATALASE HPII

Chain B:

● Molecule 1: CATALASE HPII

Chain C:
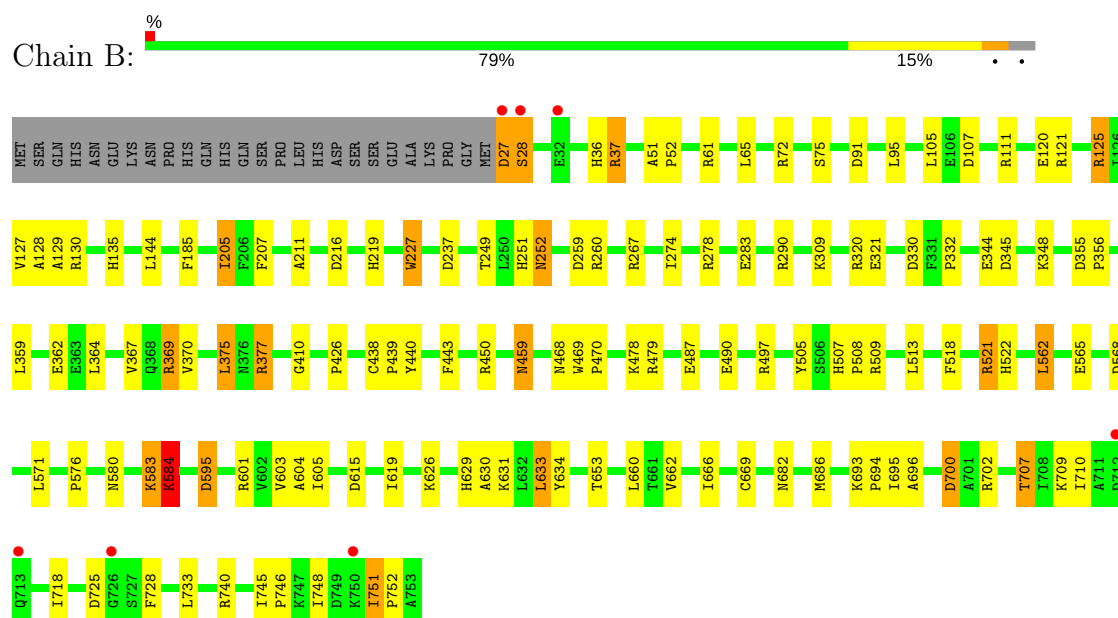
● Molecule 1: CATALASE HPII

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	93.04Å 132.34Å 121.20Å 90.00° 109.63° 90.00°	Depositor
Resolution (Å)	87.65 – 2.15 19.07 – 2.15	Depositor EDS
% Data completeness (in resolution range)	95.8 (87.65-2.15) 95.9 (19.07-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.71 (at 2.15Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.150 , 0.206 0.145 , 0.198	Depositor DCC
R_{free} test set	7176 reflections (5.26%)	DCC
Wilson B-factor (Å^2)	8.7	Xtriage
Anisotropy	0.244	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 58.6	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.026 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	25922	wwPDB-VP
Average B, all atoms (Å^2)	11.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% 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: 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	0.45	0/5896	1.13	21/8016 (0.3%)
1	B	0.45	0/5896	1.16	35/8016 (0.4%)
1	C	0.46	0/5896	1.18	31/8016 (0.4%)
1	D	0.45	0/5896	1.16	29/8016 (0.4%)
All	All	0.45	0/23584	1.16	116/32064 (0.4%)

There are no bond length outliers.

The worst 5 of 116 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	725	ASP	C-N-CA	13.45	150.54	122.30
1	C	488	ARG	CD-NE-CZ	12.38	140.93	123.60
1	D	61	ARG	NE-CZ-NH1	-11.30	114.65	120.30
1	C	121	ARG	NE-CZ-NH1	11.30	125.95	120.30
1	D	61	ARG	NE-CZ-NH2	10.25	125.43	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	5741	0	5576	74	0
1	B	5741	0	5576	77	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	5741	0	5575	81	0
1	D	5741	0	5576	78	0
2	A	43	0	30	3	0
2	B	43	0	30	1	0
2	C	43	0	30	1	0
2	D	43	0	30	3	0
3	A	737	0	0	10	0
3	B	661	0	0	9	0
3	C	675	0	0	10	0
3	D	713	0	0	7	0
All	All	25922	0	22423	280	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 6.

The worst 5 of 280 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:490:GLU:HG2	1:B:490:GLU:HG2	1.57	0.87
1:D:750:LYS:HD3	1:D:751:ILE:H	1.52	0.75
1:B:468:ASN:HD22	1:D:27:ASP:N	1.86	0.73
1:C:708:ILE:HG13	1:C:710:ILE:HG12	1.72	0.72
1:B:144:LEU:HD11	1:B:370:VAL:HG13	1.72	0.70

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	725/753 (96%)	701 (97%)	20 (3%)	4 (1%)	28	20
1	B	725/753 (96%)	700 (97%)	21 (3%)	4 (1%)	28	20

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	725/753 (96%)	701 (97%)	22 (3%)	2 (0%)	44	41
1	D	725/753 (96%)	693 (96%)	28 (4%)	4 (1%)	28	20
All	All	2900/3012 (96%)	2795 (96%)	91 (3%)	14 (0%)	32	25

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	28	SER
1	A	709	LYS
1	B	28	SER
1	C	725	ASP
1	D	28	SER

5.3.2 Protein sidechains ⓘ

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	611/635 (96%)	589 (96%)	22 (4%)	40	38
1	B	611/635 (96%)	589 (96%)	22 (4%)	40	38
1	C	611/635 (96%)	570 (93%)	41 (7%)	19	12
1	D	611/635 (96%)	580 (95%)	31 (5%)	28	22
All	All	2444/2540 (96%)	2328 (95%)	116 (5%)	30	26

5 of 116 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	205	ILE
1	C	478	LYS
1	D	616	LEU
1	C	237	ASP
1	C	283	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	449	HIS
1	C	507	HIS
1	D	546	GLN
1	C	459	ASN
1	C	572	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

4 ligands 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	#$|Z| > 2$
2	HEM	A	760	1	28,50,50	2.00	5 (17%)	17,82,82	1.37	2 (11%)
2	HEM	B	760	1	28,50,50	2.03	4 (14%)	17,82,82	1.79	5 (29%)
2	HEM	C	760	1	28,50,50	2.02	6 (21%)	17,82,82	1.88	6 (35%)
2	HEM	D	760	1	28,50,50	1.99	6 (21%)	17,82,82	1.83	4 (23%)

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	760	1	-	0/6/54/54	0/0/8/8
2	HEM	B	760	1	-	0/6/54/54	0/0/8/8
2	HEM	C	760	1	-	0/6/54/54	0/0/8/8
2	HEM	D	760	1	-	0/6/54/54	0/0/8/8

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	760	HEM	C3B-C2B	-5.18	1.33	1.40
2	B	760	HEM	C3C-C2C	-5.15	1.33	1.40
2	B	760	HEM	C3B-C2B	-5.07	1.33	1.40
2	D	760	HEM	C3B-C2B	-4.96	1.33	1.40
2	A	760	HEM	C3B-C2B	-4.94	1.33	1.40

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	760	HEM	CMD-C2D-C1D	-4.33	121.82	128.46
2	C	760	HEM	CMD-C2D-C1D	-3.50	123.08	128.46
2	C	760	HEM	CMA-C3A-C4A	-3.13	123.66	128.46
2	D	760	HEM	CMD-C2D-C1D	-3.09	123.72	128.46
2	A	760	HEM	CMA-C3A-C4A	-2.81	124.15	128.46

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
2	A	760	HEM	3	0
2	B	760	HEM	1	0
2	C	760	HEM	1	0
2	D	760	HEM	3	0

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, 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	727/753 (96%)	-0.63	7 (0%)	82	86	3, 8, 26, 50	1 (0%)
1	B	727/753 (96%)	-0.58	7 (0%)	82	86	3, 9, 28, 47	1 (0%)
1	C	727/753 (96%)	-0.59	5 (0%)	87	90	3, 9, 28, 46	1 (0%)
1	D	727/753 (96%)	-0.64	5 (0%)	87	90	3, 8, 27, 47	1 (0%)
All	All	2908/3012 (96%)	-0.61	24 (0%)	86	89	3, 8, 27, 50	4 (0%)

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	27	ASP	10.6
1	A	711	ALA	5.9
1	A	28	SER	5.7
1	A	710	ILE	5.3
1	A	712	ASP	4.5

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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, 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	LLDF	B-factors(Å^2)	Q<0.9
2	HEM	B	760	43/43	0.99	0.08	0.46	3,5,6,6	0
2	HEM	C	760	43/43	0.99	0.07	-0.33	2,5,7,8	0
2	HEM	D	760	43/43	0.99	0.07	-0.39	2,5,6,6	0
2	HEM	A	760	43/43	0.99	0.06	-0.61	2,5,6,7	0

6.5 Other polymers ⓘ

There are no such residues in this entry.