



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:25 PM GMT

PDB ID : 1TGU
Title : The crystal structure of bovine liver catalase without NADPH
Authors : Sugadev, R.; Balasundaresan, D.; Ponnuswamy, M.N.; Kumaran, D.; Swaminathan, S.; Sekar, K.
Deposited on : 2004-05-31
Resolution : 2.80 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 ⓘ 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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

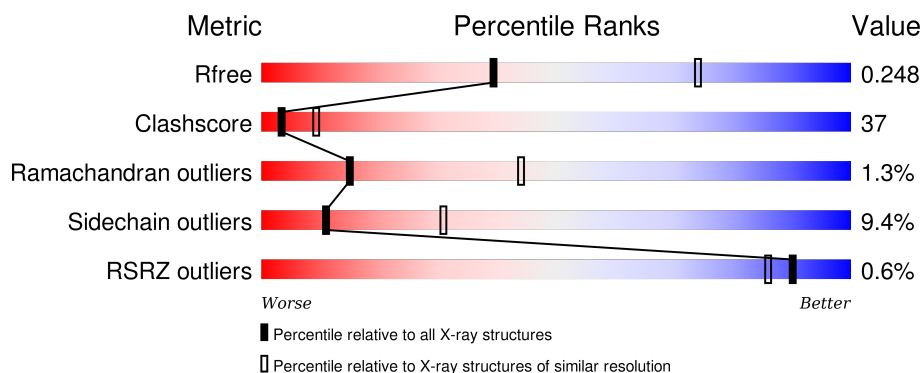
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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 $\leq 5\%$. 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	506	<div> <div>%</div> <div> <div></div> <div>54%</div> <div>38%</div> <div>6%</div> <div>..</div> </div> </div>
1	B	506	<div> <div>%</div> <div> <div></div> <div>38%</div> <div>51%</div> <div>8%</div> <div>..</div> </div> </div>
1	C	506	<div> <div></div> <div> <div>46%</div> <div>43%</div> <div>8%</div> <div>..</div> </div> </div>
1	D	506	<div> <div>%</div> <div> <div></div> <div>41%</div> <div>53%</div> <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
2	HEM	A	2000	-	-	X	-
2	HEM	B	2001	-	-	X	X
2	HEM	D	2003	-	-	X	-

2 Entry composition [i](#)

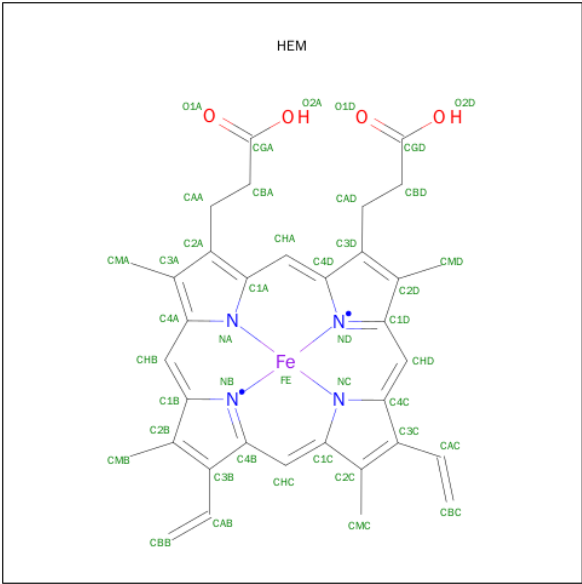
There are 3 unique types of molecules in this entry. The entry contains 17173 atoms, of which 0 are hydrogens and 0 are deuteriums.

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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	499	Total	C	N	O	S	0	0	0
			4017	2548	715	740	14			
1	B	499	Total	C	N	O	S	0	0	0
			4017	2548	715	740	14			
1	C	499	Total	C	N	O	S	1	0	0
			4017	2548	715	740	14			
1	D	499	Total	C	N	O	S	0	0	0
			4017	2548	715	740	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		

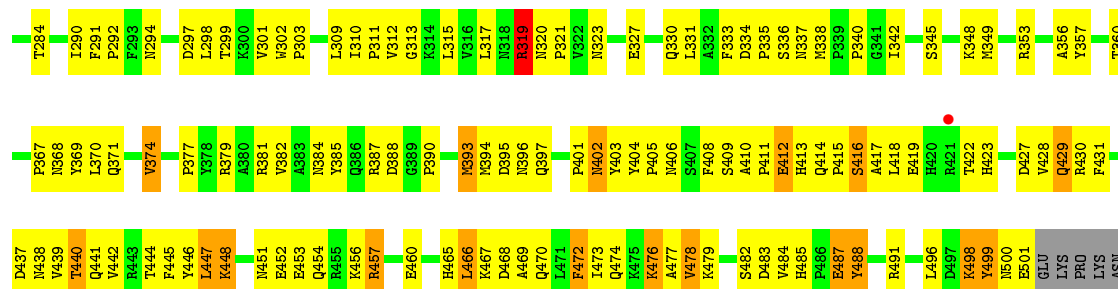
Continued on next page...

Continued from previous page...

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

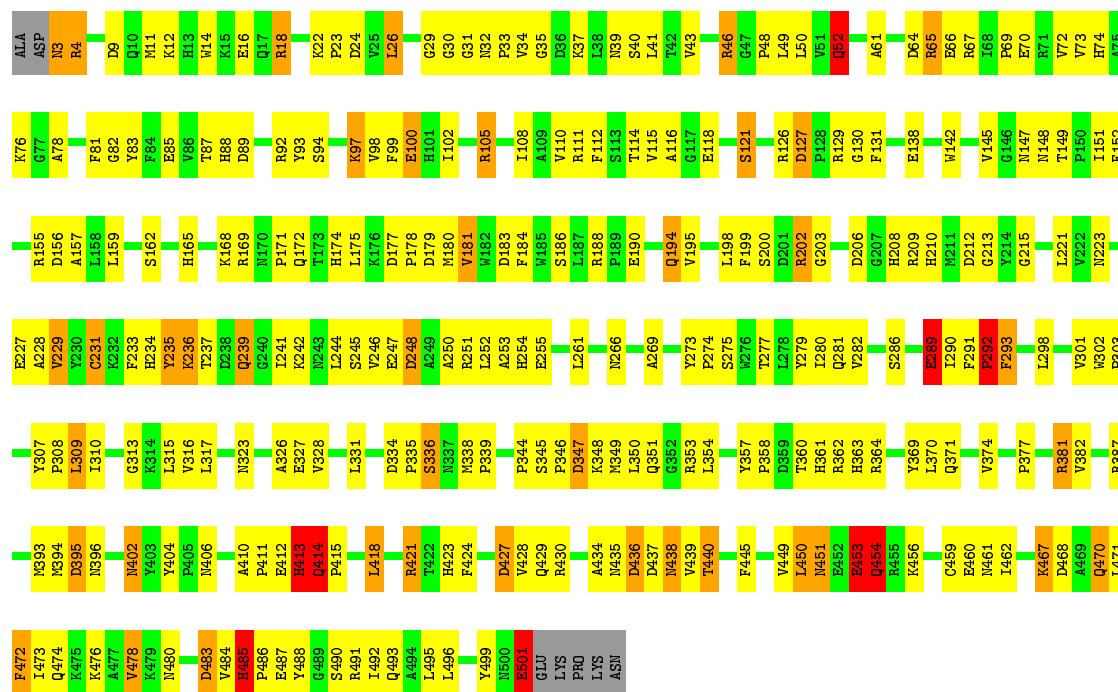
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	229	Total 229	O 229	0	0
3	B	236	Total 236	O 236	0	0
3	C	236	Total 236	O 236	0	0
3	D	232	Total 232	O 232	0	0



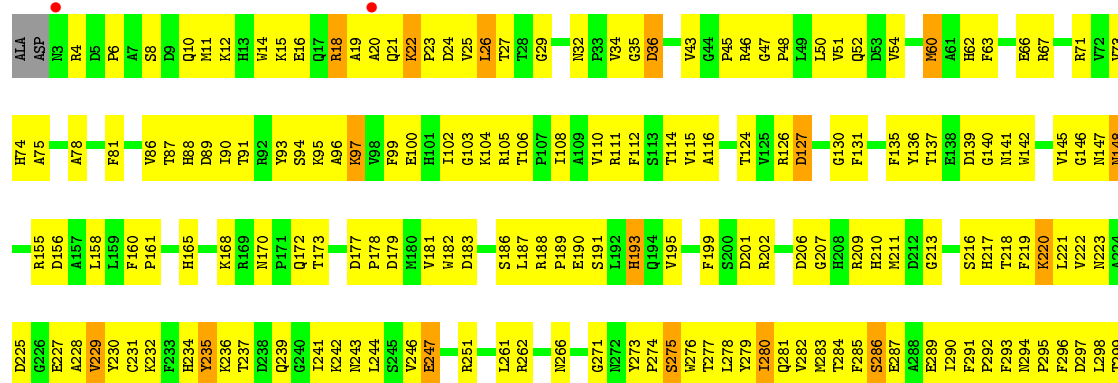
• Molecule 1: Catalase

Chain C: 46% 43% 8% ..



• Molecule 1: Catalase

Chain D: 41% 53% ..





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	85.53Å 139.67Å 225.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.80 48.76 – 2.79	Depositor EDS
% Data completeness (in resolution range)	89.9 (40.00-2.80) 89.3 (48.76-2.79)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.47 (at 2.77Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.200 , 0.249 0.199 , 0.248	Depositor DCC
R_{free} test set	1822 reflections (3.11%)	DCC
Wilson B-factor (Å ²)	29.9	Xtriage
Anisotropy	0.532	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 51.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 60796 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	17173	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.85% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 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	0.56	5/4137 (0.1%)	1.38	52/5619 (0.9%)
1	B	0.97	4/4137 (0.1%)	1.47	28/5619 (0.5%)
1	C	0.66	8/4137 (0.2%)	1.49	33/5619 (0.6%)
1	D	0.60	1/4137 (0.0%)	0.81	6/5619 (0.1%)
All	All	0.72	18/16548 (0.1%)	1.32	119/22476 (0.5%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	1
1	B	0	2
1	C	2	2
All	All	3	5

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	319	ARG	CD-NE	49.59	2.30	1.46
1	D	413	HIS	CA-CB	-20.44	1.08	1.53
1	C	202	ARG	NE-CZ	16.98	1.55	1.33
1	B	319	ARG	NE-CZ	16.07	1.53	1.33
1	C	414	GLN	CA-CB	-9.86	1.32	1.53

The worst 5 of 119 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	202	ARG	NE-CZ-NH2	-52.90	93.85	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	319	ARG	NE-CZ-NH1	-46.73	96.93	120.30
1	C	202	ARG	NE-CZ-NH1	43.98	142.29	120.30
1	B	319	ARG	CG-CD-NE	-41.22	25.24	111.80
1	B	395	ASP	N-CA-CB	-31.45	53.99	110.60

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	395	ASP	CA
1	C	453	GLU	CA
1	C	501	GLU	CA

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	421	ARG	Sidechain
1	B	126	ARG	Sidechain
1	B	319	ARG	Sidechain
1	C	289	GLU	Sidechain
1	C	292	PRO	Peptide

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	4017	0	3840	292	0
1	B	4017	0	3840	394	0
1	C	4017	0	3839	323	2
1	D	4017	0	3839	315	0
2	A	43	0	30	36	0
2	B	43	0	30	28	0
2	C	43	0	30	18	0
2	D	43	0	30	29	0
3	A	229	0	0	22	0
3	B	236	0	0	28	2
3	C	236	0	0	21	2
3	D	232	0	0	18	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	17173	0	15478	1161	3

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 37.

The worst 5 of 1161 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:421:ARG:HD2	1:B:429:GLN:CG	1.19	1.58
1:B:353:ARG:HG3	2:B:2001:HEM:CBB	1.34	1.57
1:D:147:ASN:ND2	2:D:2003:HEM:CAC	1.74	1.50
1:B:353:ARG:CG	2:B:2001:HEM:HBB2	1.02	1.49
1:D:147:ASN:ND2	2:D:2003:HEM:HAC	1.27	1.44

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:85:GLU:OE2	3:C:2084:HOH:O[3_645]	2.15	0.05
3:B:2109:HOH:O	3:C:2220:HOH:O[3_655]	2.15	0.05
1:C:102:ILE:O	3:B:2221:HOH:O[3_645]	2.16	0.04

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	497/506 (98%)	452 (91%)	42 (8%)	3 (1%)	30	65
1	B	497/506 (98%)	419 (84%)	70 (14%)	8 (2%)	12	38
1	C	497/506 (98%)	434 (87%)	55 (11%)	8 (2%)	12	38
1	D	497/506 (98%)	435 (88%)	55 (11%)	7 (1%)	14	42

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1988/2024 (98%)	1740 (88%)	222 (11%)	26 (1%)	15	44

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	19	ALA
1	B	100	GLU
1	B	124	THR
1	D	36	ASP
1	A	440	THR

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	431/437 (99%)	396 (92%)	35 (8%)	15	39
1	B	431/437 (99%)	380 (88%)	51 (12%)	6	19
1	C	431/437 (99%)	387 (90%)	44 (10%)	9	26
1	D	431/437 (99%)	399 (93%)	32 (7%)	17	43
All	All	1724/1748 (99%)	1562 (91%)	162 (9%)	11	31

5 of 162 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	402	ASN
1	C	34	VAL
1	D	392	CYS
1	B	416	SER
1	B	488	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 68 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	32	ASN
1	C	414	GLN
1	D	386	GLN
1	C	52	GLN
1	C	239	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](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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	2000	1	30,50,50	3.11	12 (40%)	24,82,82	2.80	9 (37%)
2	HEM	B	2001	1	30,50,50	3.24	13 (43%)	24,82,82	3.33	10 (41%)
2	HEM	C	2002	1	30,50,50	2.97	12 (40%)	24,82,82	2.08	8 (33%)
2	HEM	D	2003	1	30,50,50	3.52	13 (43%)	24,82,82	2.07	8 (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
2	HEM	A	2000	1	-	0/10/54/54	0/0/8/8
2	HEM	B	2001	1	-	0/10/54/54	0/0/8/8
2	HEM	C	2002	1	-	0/10/54/54	0/0/8/8
2	HEM	D	2003	1	-	0/10/54/54	0/0/8/8

The worst 5 of 50 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	2002	HEM	C3B-C4B	-6.86	1.45	1.51
2	A	2000	HEM	C3B-C4B	-6.83	1.45	1.51
2	B	2001	HEM	C3B-CAB	-6.70	1.38	1.51
2	B	2001	HEM	C3B-C4B	-6.69	1.45	1.51
2	D	2003	HEM	C3B-CAB	-6.68	1.38	1.51

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2001	HEM	C2C-C1C-CHC	-12.68	104.39	123.68
2	D	2003	HEM	CAA-C2A-C1A	-2.57	124.22	127.01
2	C	2002	HEM	CAA-C2A-C1A	-2.55	124.24	127.01
2	B	2001	HEM	CAA-C2A-C1A	-2.54	124.25	127.01
2	A	2000	HEM	CAA-C2A-C1A	-2.51	124.28	127.01

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 111 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2000	HEM	36	0
2	B	2001	HEM	28	0
2	C	2002	HEM	18	0
2	D	2003	HEM	29	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 [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	499/506 (98%)	-0.48	3 (0%) 90 86	8, 24, 48, 89	0
1	B	499/506 (98%)	-0.33	6 (1%) 81 73	10, 31, 56, 88	0
1	C	499/506 (98%)	-0.43	0 100 100	6, 27, 56, 81	1 (0%)
1	D	499/506 (98%)	-0.35	3 (0%) 90 86	10, 30, 60, 87	0
All	All	1996/2024 (98%)	-0.40	12 (0%) 90 86	6, 28, 55, 89	1 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	20	ALA	5.3
1	B	20	ALA	3.9
1	A	3	ASN	3.5
1	B	19	ALA	2.6
1	A	19	ALA	2.6

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

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, 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(\AA^2)	Q<0.9
2	HEM	B	2001	43/43	0.92	0.25	2.53	16,35,50,71	0
2	HEM	D	2003	43/43	0.93	0.22	1.76	17,33,49,60	0
2	HEM	A	2000	43/43	0.95	0.19	0.84	4,30,44,49	0
2	HEM	C	2002	43/43	0.96	0.16	0.67	15,31,42,91	0

6.5 Other polymers [i](#)

There are no such residues in this entry.