



wwPDB X-ray Structure Validation Summary Report

Jan 15, 2018 – 05:24 AM EST

PDB ID : 3EH3
Title : Structure of the reduced form of cytochrome ba3 oxidase from *Thermus thermophilus*
Authors : Liu, B.; Chen, Y.; Doukov, T.; Soltis, S.M.; Stout, D.; Fee, J.A.
Deposited on : 2008-09-11
Resolution : 3.10 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030736
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) : rb-20030736

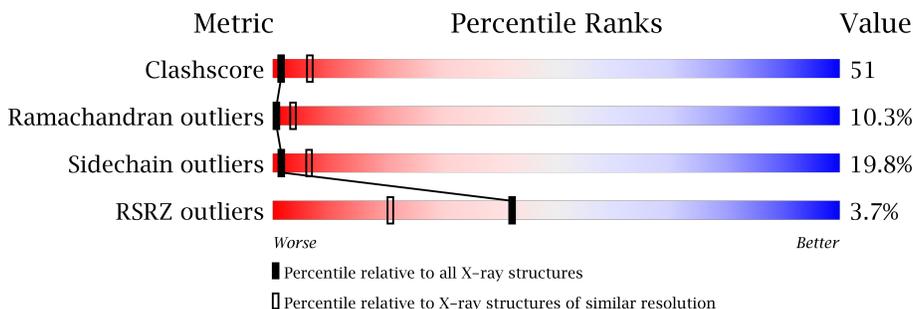
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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))
Clashscore	112137	1099 (3.12-3.08)
Ramachandran outliers	110173	1057 (3.12-3.08)
Sidechain outliers	110143	1057 (3.12-3.08)
RSRZ outliers	101464	1006 (3.12-3.08)

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	618	 4% 27% 45% 16% 10%
2	B	166	 2% 35% 42% 21%
3	C	33	 6% 39% 33% 21% 6%

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	HEM	A	800	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	HAS	A	801	-	-	X	-

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 6077 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 Cytochrome c oxidase subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	557	4409	2985	709	699	16	0	0	0

There are 58 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-55	SER	-	EXPRESSION TAG	UNP Q5SJ79
A	-54	GLU	-	EXPRESSION TAG	UNP Q5SJ79
A	-53	ILE	-	EXPRESSION TAG	UNP Q5SJ79
A	-52	SER	-	EXPRESSION TAG	UNP Q5SJ79
A	-51	ARG	-	EXPRESSION TAG	UNP Q5SJ79
A	-50	VAL	-	EXPRESSION TAG	UNP Q5SJ79
A	-49	TYR	-	EXPRESSION TAG	UNP Q5SJ79
A	-48	GLU	-	EXPRESSION TAG	UNP Q5SJ79
A	-47	ALA	-	EXPRESSION TAG	UNP Q5SJ79
A	-46	TYR	-	EXPRESSION TAG	UNP Q5SJ79
A	-45	PRO	-	EXPRESSION TAG	UNP Q5SJ79
A	-44	GLU	-	EXPRESSION TAG	UNP Q5SJ79
A	-43	LYS	-	EXPRESSION TAG	UNP Q5SJ79
A	-42	LYS	-	EXPRESSION TAG	UNP Q5SJ79
A	-41	ALA	-	EXPRESSION TAG	UNP Q5SJ79
A	-40	THR	-	EXPRESSION TAG	UNP Q5SJ79
A	-39	LEU	-	EXPRESSION TAG	UNP Q5SJ79
A	-38	TYR	-	EXPRESSION TAG	UNP Q5SJ79
A	-37	PHE	-	EXPRESSION TAG	UNP Q5SJ79
A	-36	LEU	-	EXPRESSION TAG	UNP Q5SJ79
A	-35	VAL	-	EXPRESSION TAG	UNP Q5SJ79
A	-34	LEU	-	EXPRESSION TAG	UNP Q5SJ79
A	-33	GLY	-	EXPRESSION TAG	UNP Q5SJ79
A	-32	PHE	-	EXPRESSION TAG	UNP Q5SJ79
A	-31	LEU	-	EXPRESSION TAG	UNP Q5SJ79
A	-30	ALA	-	EXPRESSION TAG	UNP Q5SJ79
A	-29	LEU	-	EXPRESSION TAG	UNP Q5SJ79

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-28	ILE	-	EXPRESSION TAG	UNP Q5SJ79
A	-27	VAL	-	EXPRESSION TAG	UNP Q5SJ79
A	-26	GLY	-	EXPRESSION TAG	UNP Q5SJ79
A	-25	SER	-	EXPRESSION TAG	UNP Q5SJ79
A	-24	LEU	-	EXPRESSION TAG	UNP Q5SJ79
A	-23	PHE	-	EXPRESSION TAG	UNP Q5SJ79
A	-22	GLY	-	EXPRESSION TAG	UNP Q5SJ79
A	-21	PRO	-	EXPRESSION TAG	UNP Q5SJ79
A	-20	PHE	-	EXPRESSION TAG	UNP Q5SJ79
A	-19	GLN	-	EXPRESSION TAG	UNP Q5SJ79
A	-18	ALA	-	EXPRESSION TAG	UNP Q5SJ79
A	-17	LEU	-	EXPRESSION TAG	UNP Q5SJ79
A	-16	ASN	-	EXPRESSION TAG	UNP Q5SJ79
A	-15	TYR	-	EXPRESSION TAG	UNP Q5SJ79
A	-14	GLY	-	EXPRESSION TAG	UNP Q5SJ79
A	-13	ASN	-	EXPRESSION TAG	UNP Q5SJ79
A	-12	VAL	-	EXPRESSION TAG	UNP Q5SJ79
A	-11	ASP	-	EXPRESSION TAG	UNP Q5SJ79
A	-10	ALA	-	EXPRESSION TAG	UNP Q5SJ79
A	-9	TYR	-	EXPRESSION TAG	UNP Q5SJ79
A	-8	PRO	-	EXPRESSION TAG	UNP Q5SJ79
A	-7	LEU	-	EXPRESSION TAG	UNP Q5SJ79
A	-6	LEU	-	EXPRESSION TAG	UNP Q5SJ79
A	-5	MET	-	EXPRESSION TAG	UNP Q5SJ79
A	-4	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	-3	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	-2	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	-1	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	0	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	1	HIS	-	EXPRESSION TAG	UNP Q5SJ79
A	258	ARG	LYS	ENGINEERED	UNP Q5SJ79

- Molecule 2 is a protein called Cytochrome c oxidase subunit 2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O	S			
2	B	166	1298	844	217	233	4	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	4	GLN	GLU	ENGINEERED	UNP Q5SJ80

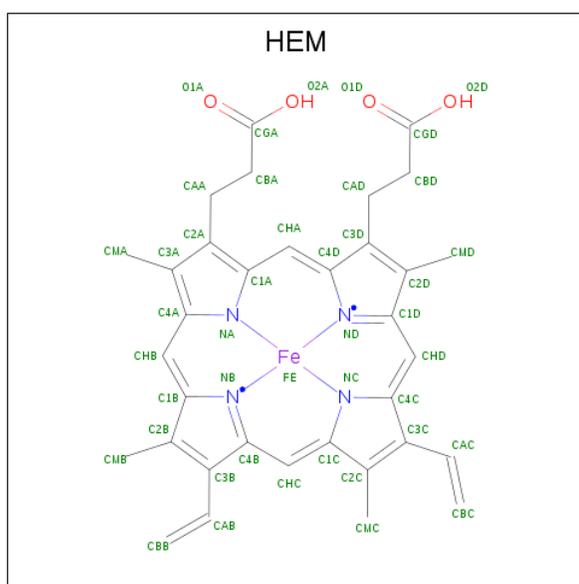
- Molecule 3 is a protein called Cytochrome c oxidase polypeptide 2A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	33	259	179	39	41	0	0	0

- Molecule 4 is COPPER (I) ION (three-letter code: CU1) (formula: Cu).

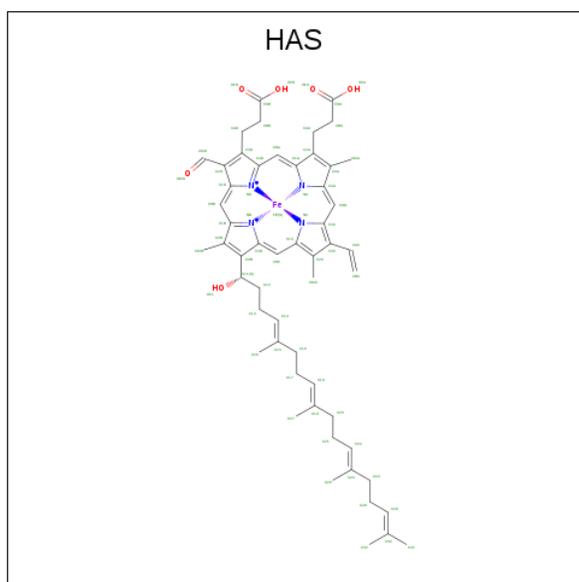
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cu		
4	A	1	1	1	0	0

- Molecule 5 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).



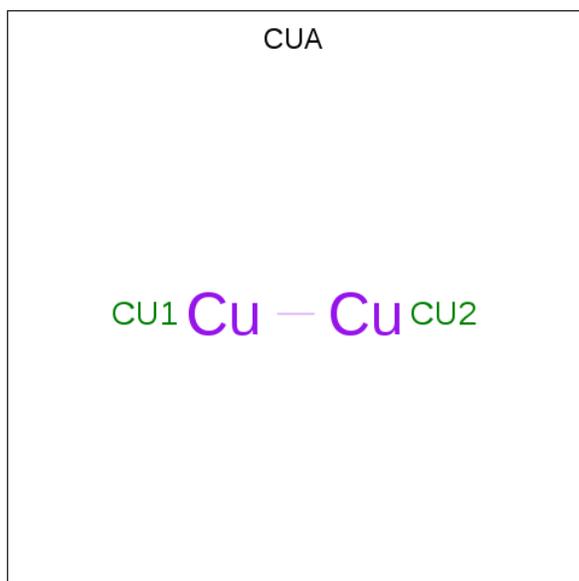
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Fe	N	O		
5	A	1	43	34	1	4	4	0	0

- Molecule 6 is HEME-AS (three-letter code: HAS) (formula: $C_{54}H_{64}FeN_4O_6$).

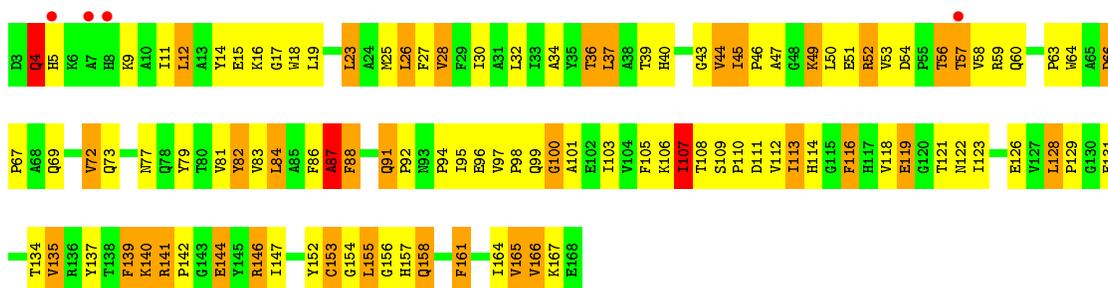


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	Fe	N			O
6	A	1	65	54	1	4	6	0	0

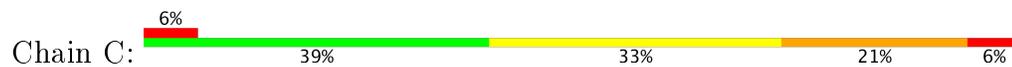
- Molecule 7 is DINUCLEAR COPPER ION (three-letter code: CUA) (formula: Cu₂).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cu		
7	B	1	2	2	0	0



- Molecule 3: Cytochrome c oxidase polypeptide 2A



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	119.75Å 119.75Å 150.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 3.10 19.99 – 2.95	Depositor EDS
% Data completeness (in resolution range)	96.0 (19.99-3.10) 92.4 (19.99-2.95)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.03 (at 2.93Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.201 , 0.300 0.209 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	115.0	Xtrriage
Anisotropy	0.092	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 129.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6077	wwPDB-VP
Average B, all atoms (Å ²)	101.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.06% 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, 0.333 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: HEM, HAS, CUA, CU1

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	1.15	10/4566 (0.2%)	1.25	27/6266 (0.4%)
2	B	1.10	0/1335	1.20	8/1822 (0.4%)
3	C	1.20	0/265	1.31	2/359 (0.6%)
All	All	1.14	10/6166 (0.2%)	1.24	37/8447 (0.4%)

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	0	2
2	B	0	1
3	C	0	1
All	All	0	4

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	52	TYR	CE1-CZ	5.69	1.46	1.38
1	A	52	TYR	CD1-CE1	5.62	1.47	1.39
1	A	236	VAL	CB-CG2	5.52	1.64	1.52
1	A	230	TRP	CB-CG	5.43	1.60	1.50
1	A	56	LYS	CD-CE	5.42	1.64	1.51

The worst 5 of 37 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	59	LEU	CA-CB-CG	9.98	138.26	115.30
1	A	387	LEU	CA-CB-CG	-9.97	92.37	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	20	LEU	CA-CB-CG	9.85	137.96	115.30
1	A	241	LEU	CB-CG-CD2	-9.82	94.31	111.00
1	A	449	ARG	NE-CZ-NH2	-8.64	115.98	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	137	PRO	Peptide
1	A	8	ILE	Peptide
2	B	87	ALA	Peptide
3	C	33	ARG	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	4409	0	4516	507	0
2	B	1298	0	1282	140	0
3	C	259	0	279	30	0
4	A	1	0	0	0	0
5	A	43	0	30	25	0
6	A	65	0	61	21	0
7	B	2	0	0	0	0
All	All	6077	0	6168	623	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 51.

The worst 5 of 623 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:233:HIS:NE2	1:A:237:TYR:HE2	0.84	1.31
1:A:233:HIS:NE2	1:A:237:TYR:CE2	1.77	1.17
1:A:101:PRO:HA	1:A:166:LEU:HD21	1.15	1.15

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:67:PRO:HB3	2:B:91:GLN:HG3	1.26	1.15
2:B:50:LEU:HD13	2:B:52:ARG:HE	1.17	1.08

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	555/618 (90%)	385 (69%)	108 (20%)	62 (11%)	0	2
2	B	164/166 (99%)	131 (80%)	23 (14%)	10 (6%)	2	11
3	C	31/33 (94%)	24 (77%)	2 (6%)	5 (16%)	0	0
All	All	750/817 (92%)	540 (72%)	133 (18%)	77 (10%)	0	3

5 of 77 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	11	VAL
1	A	36	SER
1	A	102	ASN
1	A	127	ASN
1	A	161	TYR

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	453/503 (90%)	376 (83%)	77 (17%)	2	11
2	B	136/136 (100%)	98 (72%)	38 (28%)	0	1
3	C	26/26 (100%)	19 (73%)	7 (27%)	0	1
All	All	615/665 (92%)	493 (80%)	122 (20%)	1	6

5 of 122 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	470	ASN
1	A	524	MET
2	B	165	VAL
1	A	472	LEU
1	A	492	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	446	ASN
2	B	73	GLN
2	B	114	HIS
1	A	440	HIS
2	B	91	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](#)

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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	HEM	A	800	1	28,50,50	3.60	13 (46%)	17,82,82	3.05	8 (47%)
6	HAS	A	801	1	55,72,72	4.23	21 (38%)	48,109,109	3.36	26 (54%)
7	CUA	B	802	2	0,1,1	0.00	-	0,0,0	0.00	-

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	HEM	A	800	1	-	0/6/54/54	0/0/8/8
6	HAS	A	801	1	-	0/35/122/122	0/0/8/8
7	CUA	B	802	2	-	0/0/0/0	0/0/0/0

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	801	HAS	C1B-NB	-13.44	1.31	1.49
6	A	801	HAS	C1D-ND	-11.50	1.34	1.49
6	A	801	HAS	C4D-ND	-10.21	1.35	1.49
6	A	801	HAS	C4B-NB	-8.81	1.37	1.49
6	A	801	HAS	C3C-CAC	-7.62	1.32	1.47

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	800	HEM	CAA-CBA-CGA	-7.33	100.13	112.66
6	A	801	HAS	CAD-CBD-CGD	-6.69	101.23	112.66
6	A	801	HAS	CAA-CBA-CGA	-6.66	101.27	112.66
6	A	801	HAS	CHC-C1C-C2C	-5.37	120.16	129.45
6	A	801	HAS	C4A-C3A-C2A	-4.50	100.39	105.82

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 46 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	800	HEM	25	0
6	A	801	HAS	21	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 [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	557/618 (90%)	-0.02	22 (3%) 40 19	66, 98, 130, 159	0
2	B	166/166 (100%)	0.05	4 (2%) 59 37	75, 98, 134, 192	0
3	C	33/33 (100%)	-0.26	2 (6%) 22 9	73, 88, 128, 146	0
All	All	756/817 (92%)	-0.01	28 (3%) 42 21	66, 98, 133, 192	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	6	SER	15.1
1	A	7	GLU	12.0
1	A	493	LEU	6.0
2	B	5	HIS	4.6
1	A	14	ALA	4.6

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(Å ²)	Q<0.9
6	HAS	A	801	65/65	0.96	0.19	-0.11	65,82,93,95	0
5	HEM	A	800	43/43	0.97	0.18	-0.34	66,79,100,118	0
7	CUA	B	802	2/2	0.99	0.09	-1.92	103,103,103,118	0
4	CU1	A	803	1/1	0.98	0.15	-	97,97,97,97	0

6.5 Other polymers [i](#)

There are no such residues in this entry.