wwPDB X-ray Structure Validation Summary Report (i)

Aug 4, 2026 – 09:54 PM EDT

PDB ID	:	2QPD / pdb_00002qpd
Title	:	An unexpected outcome of surface-engineering an integral membrane protein: Improved crystallization of cytochrome ba3 oxidase from Thermus thermophilus
Authors	:	Liu, B.; Luna, V.M.; Chen, Y.; Stout, C.D.; Fee, J.A.
Deposited on	:	2007-07-23
Resolution	:	3.25 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

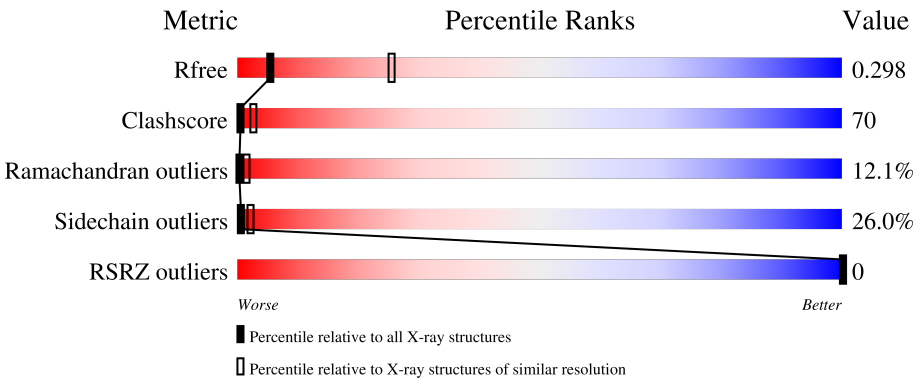
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.25 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}	180053	1605 (3.30-3.22)
Clashscore	190562	1660 (3.30-3.22)
Ramachandran outliers	187476	1630 (3.30-3.22)
Sidechain outliers	187428	1629 (3.30-3.22)
RSRZ outliers	180081	1605 (3.30-3.22)

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 >=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 <=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	568	17% 45% 31% 5% •
2	B	168	17% 49% 27% 5% •
3	C	34	9% 47% 35% 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
6	HAS	A	801	X	-	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 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
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 mutation	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	216	234	4	0	0	0

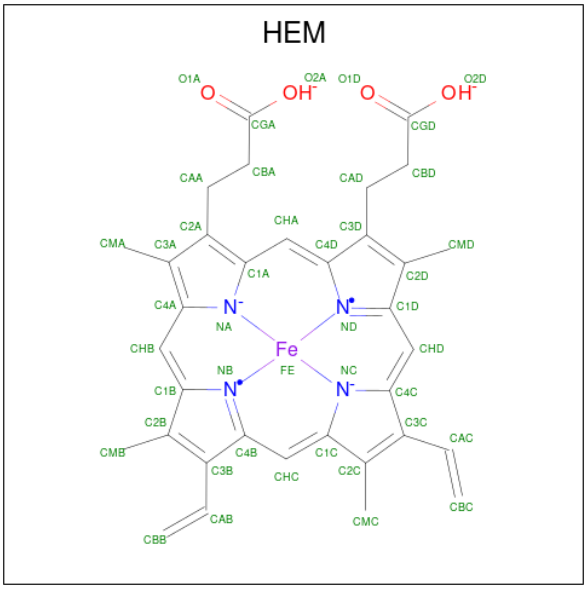
- Molecule 3 is a protein called Cytochrome c oxidase polypeptide 2A.

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 (CCD ID: CU1) (formula: Cu).

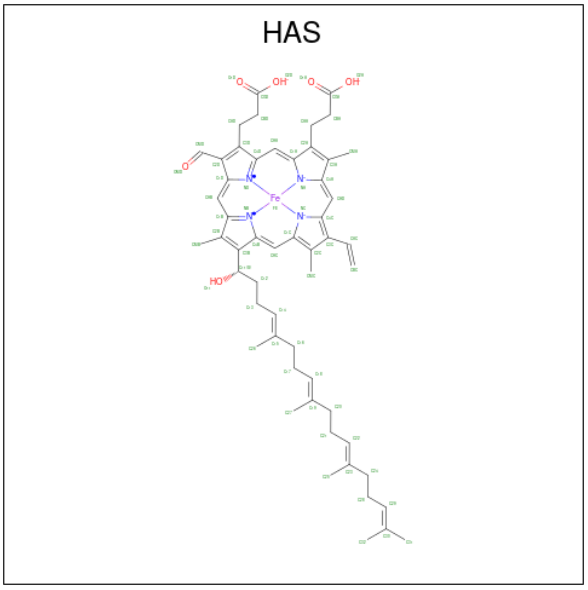
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cu		
4	A	1	1	1	0	0

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



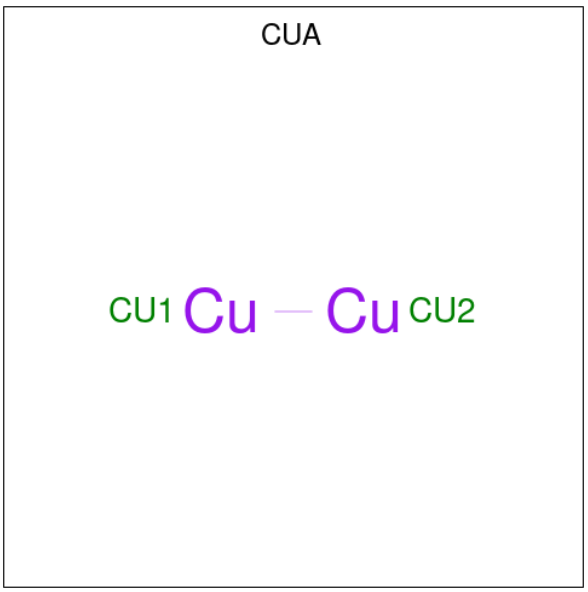
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Fe	N	O		
5	A	1	43	34	1	4	4	0	0

- Molecule 6 is HEME-AS (CCD ID: 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 (CCD ID: CUA) (formula: Cu_2).



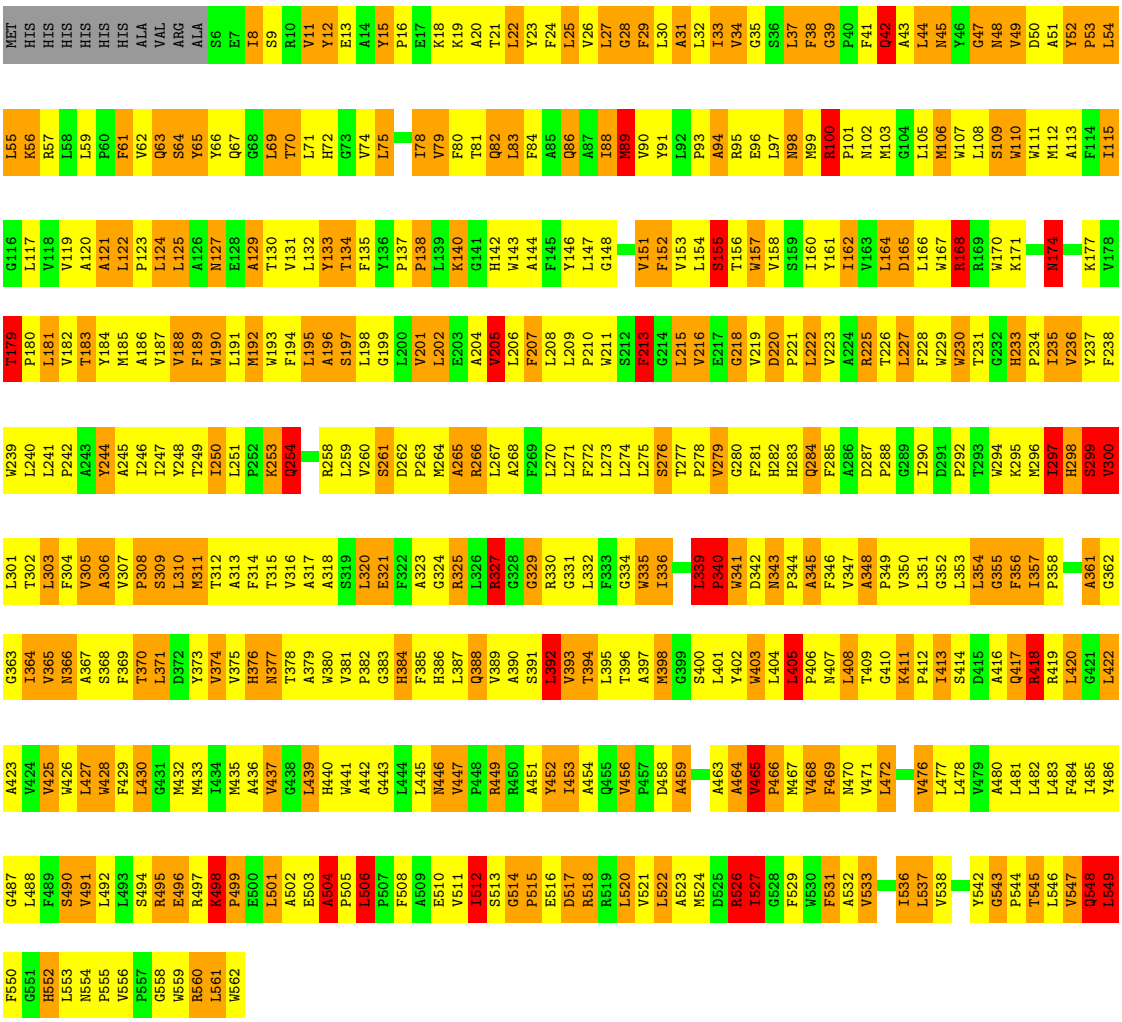
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cu		
7	B	1	2	2	0	0

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: Cytochrome c oxidase subunit 1

Chain A: 17% 45% 31% 5% •



- Molecule 2: Cytochrome c oxidase subunit 2

Chain B: 17% 49% 27% 5% •

MET VAL D3 K6 A7 H8 K9 A10 I11 L12 A13 Y14 E15 K16 G17 W18 L19 S22 L23 A24 M25 L26 F27 V28 F29 I30 A31 L32 I33 A34 Y35 T36 L37 A38 T39 H40 T41 V44 I45 P46 A47 G48 K49 L50 E51 R52 V53 D54 P55 T56 T57 V58 R59 Q60 E61 G62 P63

W64 A65 D66 P67 A68 Q69 A70 V71 V72 Q73 T74 G75 P76 N77 Q78 Y79 T80 V81 Y82 V83 L84 A85 F86 A87 F88 G89 Y90 Q91 P92 N93 P94 I95 E96 V97 P98 Q99 G100 A101 E102 I103 V104 I107 T108 S109 P110 D111 V112 I113 H114 G115 F116 H117 V118 E119 G120 T121 N122 I123 N124

V125 E126 V127 L128 P129 G130 E131 V132 S133 T134 V135 R136 Y137 T138 F139 K140 R141 P142 G143 E144 Y145 R146 I147 I148 C149 N150 Q151 Y152 C153 G154 L155 G156 H157 Q158 N159 M160 F161 G162 T163 I164 V165 V166 K167 E168

- Molecule 3: Cytochrome c oxidase polypeptide 2A

Chain C: 9% 47% 35% 6%

MET E2 E3 K4 P5 K6 G7 A8 L9 A10 V11 I12 L13 V14 L15 T16 L17 T18 I19 L20 V21 F22 W23 L24 G25 V26 Y27 A28 V29 F30 F31 A32 R33 G34

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	115.19Å 115.19Å 149.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.98 – 3.25 19.98 – 3.25	Depositor EDS
% Data completeness (in resolution range)	77.2 (19.98-3.25) 82.2 (19.98-3.25)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.25 (at 3.04Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.215 , 0.307 0.189 , 0.298	Depositor DCC
R_{free} test set	780 reflections (3.82%)	wwPDB-VP
Wilson B-factor (Å^2)	114.7	Xtriage
Anisotropy	0.167	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 71.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6077	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage'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.*

[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: CUA, HEM, CU1, HAS

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	1.32	21/4566 (0.5%)	1.60	83/6266 (1.3%)
2	B	1.38	12/1335 (0.9%)	1.67	30/1822 (1.6%)
3	C	1.43	2/265 (0.8%)	1.45	3/359 (0.8%)
All	All	1.34	35/6166 (0.6%)	1.61	116/8447 (1.4%)

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	4
2	B	0	3
All	All	0	7

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	72	VAL	CA-CB	7.35	1.64	1.54
1	A	447	VAL	CA-CB	6.67	1.62	1.54
2	B	73	GLN	CA-C	6.49	1.61	1.52
2	B	52	ARG	CA-C	6.31	1.60	1.52
1	A	205	VAL	CA-CB	6.28	1.60	1.54

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	339	LEU	CA-C-N	11.37	134.06	119.84
1	A	339	LEU	C-N-CA	11.37	134.06	119.84
1	A	44	LEU	N-CA-C	-10.52	100.58	113.41

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	514	GLY	CA-C-N	10.36	132.79	119.84
1	A	514	GLY	C-N-CA	10.36	132.79	119.84

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	174	ASN	Peptide
1	A	400	SER	Peptide
1	A	458	ASP	Peptide
1	A	503	GLU	Peptide
2	B	46	PRO	Peptide

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	4409	0	4516	684	0
2	B	1298	0	1280	166	0
3	C	259	0	279	45	0
4	A	1	0	0	0	0
5	A	43	0	30	9	0
6	A	65	0	61	31	0
7	B	2	0	0	0	0
All	All	6077	0	6166	858	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 70.

The worst 5 of 858 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	1.02	1.42
1:A:233:HIS:NE2	1:A:237:TYR:CE2	1.95	1.27
1:A:12:TYR:HA	1:A:16:PRO:HB3	1.21	1.20

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:102:ASN:OD1	1:A:105:LEU:HG	1.41	1.19
1:A:120:ALA:O	1:A:123:PRO:HD2	1.41	1.17

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	555/568 (98%)	359 (65%)	126 (23%)	70 (13%)	0 1
2	B	164/168 (98%)	109 (66%)	40 (24%)	15 (9%)	0 3
3	C	31/34 (91%)	15 (48%)	10 (32%)	6 (19%)	0 0
All	All	750/770 (97%)	483 (64%)	176 (24%)	91 (12%)	0 1

5 of 91 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	11	VAL
1	A	12	TYR
1	A	48	ASN
1	A	151	VAL
1	A	152	PHE

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	453/462 (98%)	338 (75%)	115 (25%)	0 2
2	B	136/138 (99%)	99 (73%)	37 (27%)	0 2
3	C	26/27 (96%)	18 (69%)	8 (31%)	0 1
All	All	615/627 (98%)	455 (74%)	160 (26%)	0 2

5 of 160 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	9	LYS
2	B	140	LYS
2	B	18	TRP
2	B	72	VAL
2	B	166	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	117	HIS
2	B	159	ASN
1	A	470	ASN
1	A	554	ASN
2	B	8	HIS

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 oligosaccharides 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	$\#\|Z\| > 2$
6	HAS	A	801	1	72,72,72	2.82	31 (43%)	87,109,109	4.43	45 (51%)
5	HEM	A	800	1	50,50,50	2.14	13 (26%)	67,82,82	1.83	16 (23%)
7	CUA	B	802	2	0,1,1	-	-	-		

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
6	HAS	A	801	1	1/1/8/18	12/42/82/82	-
5	HEM	A	800	1	-	6/14/54/54	-

The worst 5 of 44 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	800	HEM	C3D-C2D	7.71	1.53	1.36
6	A	801	HAS	C4A-NA	-7.65	1.25	1.39
6	A	801	HAS	C1A-NA	-6.50	1.27	1.39
6	A	801	HAS	C2A-C3A	6.23	1.50	1.36
6	A	801	HAS	CHC-C4B	5.92	1.50	1.38

The worst 5 of 61 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	801	HAS	C3C-C2C-C1C	-14.49	90.09	107.17
6	A	801	HAS	CHA-C1A-C2A	-12.12	105.73	124.86
6	A	801	HAS	C2D-C3D-C4D	-11.90	97.82	106.43
6	A	801	HAS	C2A-C1A-NA	11.37	121.29	110.32
6	A	801	HAS	CBC-CAC-C3C	9.07	172.82	127.53

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	A	801	HAS	NA

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	800	HEM	C2B-C3B-CAB-CBB
5	A	800	HEM	C2C-C3C-CAC-CBC
5	A	800	HEM	C4C-C3C-CAC-CBC
6	A	801	HAS	C1D-C2D-CMD-OMD
6	A	801	HAS	C3D-C2D-CMD-OMD

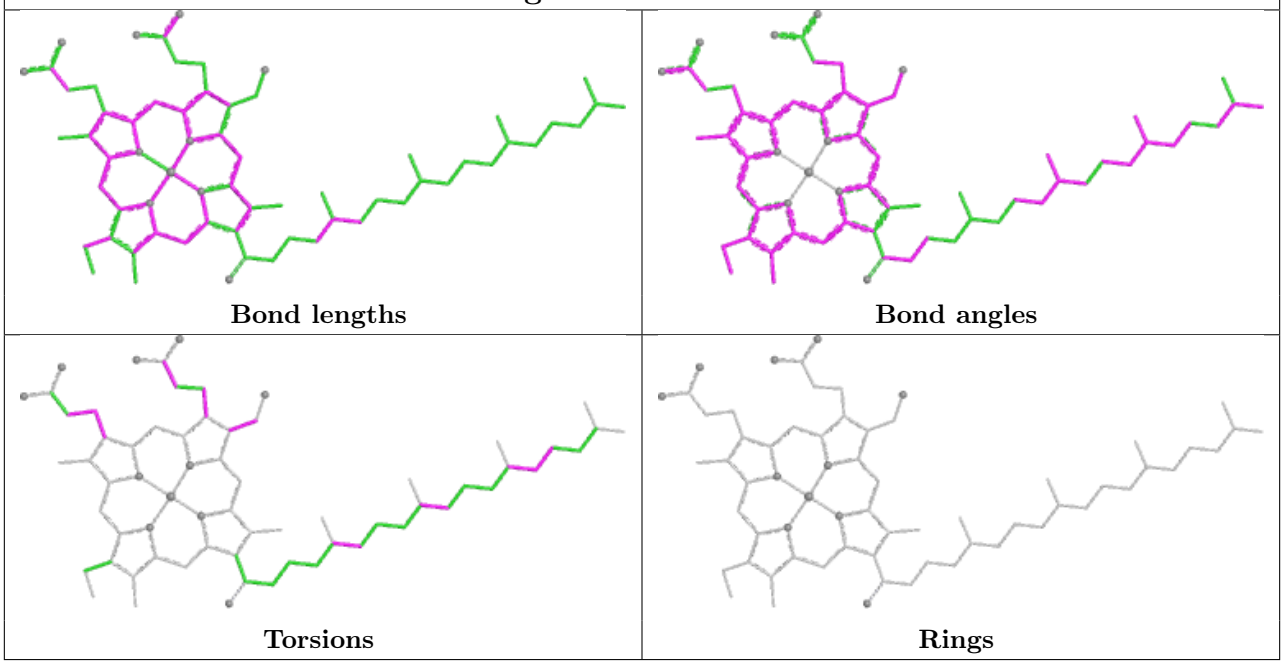
There are no ring outliers.

2 monomers are involved in 40 short contacts:

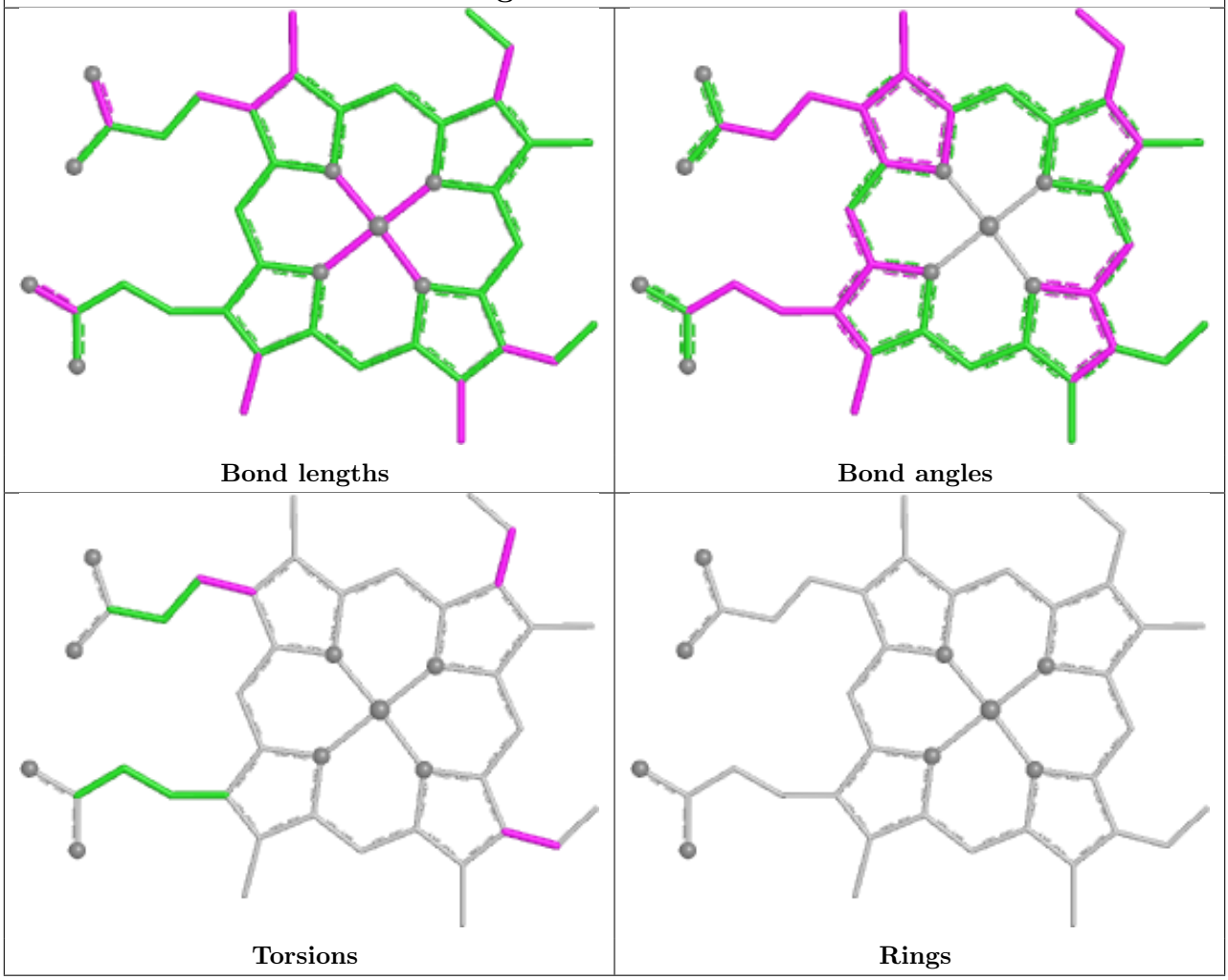
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	801	HAS	31	0
5	A	800	HEM	9	0

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 then 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.

Ligand HAS A 801



Ligand HEM A 800



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	557/568 (98%)	-0.84	0	100	100	8, 39, 82, 189	0
2	B	166/168 (98%)	-0.96	0	100	100	11, 39, 66, 150	0
3	C	33/34 (97%)	-0.87	0	100	100	16, 28, 99, 128	0
All	All	756/770 (98%)	-0.86	0	100	100	8, 39, 82, 189	0

There are no RSRZ outliers to report.

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 oligosaccharides 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, 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.

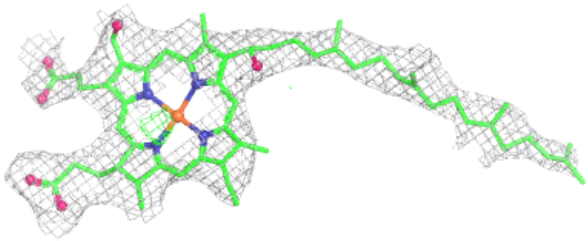
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	HAS	A	801	65/65	0.96	0.08	17,54,76,85	0
5	HEM	A	800	43/43	0.97	0.08	44,71,84,86	0
4	CU1	A	803	1/1	0.97	0.03	83,83,83,83	0
7	CUA	B	802	2/2	0.99	0.03	49,49,49,52	0

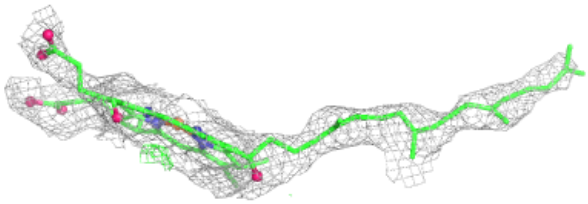
The following is a graphical depiction of the model fit to experimental electron density of all

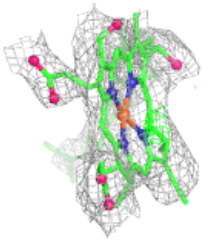
instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around HAS A 801:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

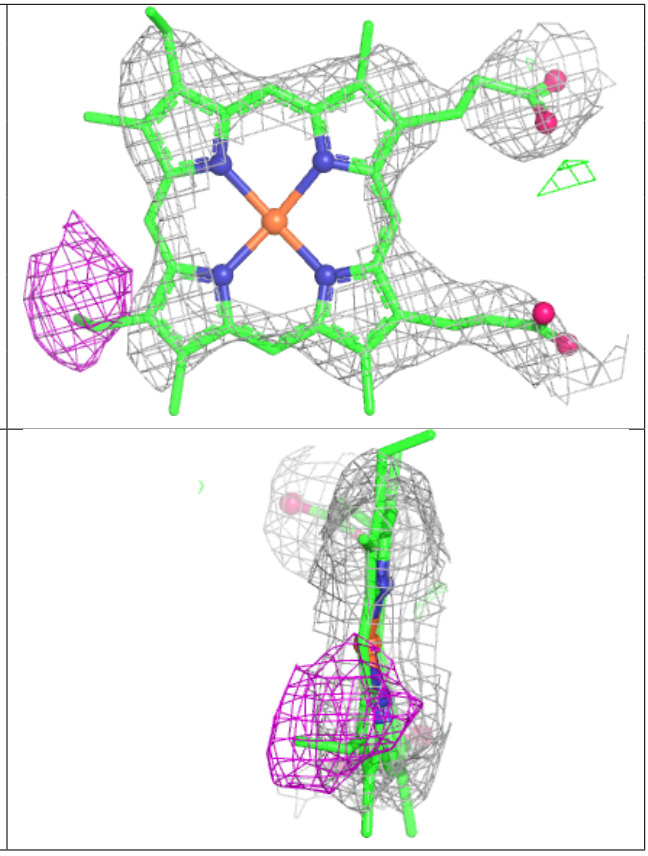






Electron density around HEM A 800:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.