



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 07:12 AM GMT

PDB ID : 4L27  
Title : Crystal structure of delta1-39 and delta516-525 human cystathionine beta-synthase D444N mutant containing C-terminal 6xHis tag  
Authors : Ereno, J.; Majtan, T.; Oyenarte, I.; Kraus, J.P.; Martinez, L.A.  
Deposited on : 2013-06-04  
Resolution : 3.39 Å(reported)

This is a wwPDB validation summary 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/ValidationPDFNotes.html>

---

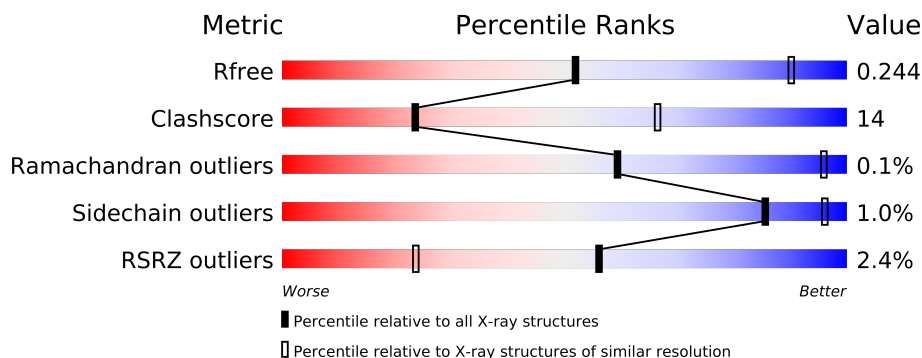
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.39 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))
$R_{free}$	66092	1017 (3.52-3.28)
Clashscore	79885	1214 (3.50-3.30)
Ramachandran outliers	78287	1177 (3.50-3.30)
Sidechain outliers	78261	1177 (3.50-3.30)
RSRZ outliers	66119	1017 (3.52-3.28)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	548	
1	B	548	
1	C	548	
1	D	548	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	HEM	C	602	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15361 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Cystathionine beta-synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	490	Total	C	N	O	S	0	0	0
			3791	2402	661	706	22			
1	D	491	Total	C	N	O	S	0	0	0
			3788	2398	662	706	22			
1	A	485	Total	C	N	O	S	0	0	0
			3757	2381	656	698	22			
1	C	491	Total	C	N	O	S	0	0	0
			3793	2403	665	703	22			

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	444	ASN	ASP	ENGINEERED MUTATION	UNP P35520
B	?	-	ILE	DELETION	UNP P35520
B	?	-	GLN	DELETION	UNP P35520
B	?	-	TYR	DELETION	UNP P35520
B	?	-	HIS	DELETION	UNP P35520
B	?	-	SER	DELETION	UNP P35520
B	?	-	THR	DELETION	UNP P35520
B	?	-	GLY	DELETION	UNP P35520
B	?	-	LYS	DELETION	UNP P35520
B	?	-	SER	DELETION	UNP P35520
B	?	-	SER	DELETION	UNP P35520
B	552	LEU	-	EXPRESSION TAG	UNP P35520
B	553	GLU	-	EXPRESSION TAG	UNP P35520
B	554	HIS	-	EXPRESSION TAG	UNP P35520
B	555	HIS	-	EXPRESSION TAG	UNP P35520
B	556	HIS	-	EXPRESSION TAG	UNP P35520
B	557	HIS	-	EXPRESSION TAG	UNP P35520
B	558	HIS	-	EXPRESSION TAG	UNP P35520
B	559	HIS	-	EXPRESSION TAG	UNP P35520
D	444	ASN	ASP	ENGINEERED MUTATION	UNP P35520
D	?	-	ILE	DELETION	UNP P35520

*Continued on next page...*

*Continued from previous page...*

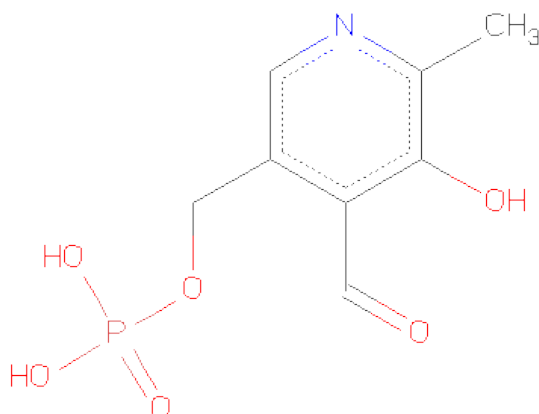
Chain	Residue	Modelled	Actual	Comment	Reference
D	?	-	GLN	DELETION	UNP P35520
D	?	-	TYR	DELETION	UNP P35520
D	?	-	HIS	DELETION	UNP P35520
D	?	-	SER	DELETION	UNP P35520
D	?	-	THR	DELETION	UNP P35520
D	?	-	GLY	DELETION	UNP P35520
D	?	-	LYS	DELETION	UNP P35520
D	?	-	SER	DELETION	UNP P35520
D	?	-	SER	DELETION	UNP P35520
D	552	LEU	-	EXPRESSION TAG	UNP P35520
D	553	GLU	-	EXPRESSION TAG	UNP P35520
D	554	HIS	-	EXPRESSION TAG	UNP P35520
D	555	HIS	-	EXPRESSION TAG	UNP P35520
D	556	HIS	-	EXPRESSION TAG	UNP P35520
D	557	HIS	-	EXPRESSION TAG	UNP P35520
D	558	HIS	-	EXPRESSION TAG	UNP P35520
D	559	HIS	-	EXPRESSION TAG	UNP P35520
A	444	ASN	ASP	ENGINEERED MUTATION	UNP P35520
A	?	-	ILE	DELETION	UNP P35520
A	?	-	GLN	DELETION	UNP P35520
A	?	-	TYR	DELETION	UNP P35520
A	?	-	HIS	DELETION	UNP P35520
A	?	-	SER	DELETION	UNP P35520
A	?	-	THR	DELETION	UNP P35520
A	?	-	GLY	DELETION	UNP P35520
A	?	-	LYS	DELETION	UNP P35520
A	?	-	SER	DELETION	UNP P35520
A	?	-	SER	DELETION	UNP P35520
A	552	LEU	-	EXPRESSION TAG	UNP P35520
A	553	GLU	-	EXPRESSION TAG	UNP P35520
A	554	HIS	-	EXPRESSION TAG	UNP P35520
A	555	HIS	-	EXPRESSION TAG	UNP P35520
A	556	HIS	-	EXPRESSION TAG	UNP P35520
A	557	HIS	-	EXPRESSION TAG	UNP P35520
A	558	HIS	-	EXPRESSION TAG	UNP P35520
A	559	HIS	-	EXPRESSION TAG	UNP P35520
C	444	ASN	ASP	ENGINEERED MUTATION	UNP P35520
C	?	-	ILE	DELETION	UNP P35520
C	?	-	GLN	DELETION	UNP P35520
C	?	-	TYR	DELETION	UNP P35520
C	?	-	HIS	DELETION	UNP P35520
C	?	-	SER	DELETION	UNP P35520

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	THR	DELETION	UNP P35520
C	?	-	GLY	DELETION	UNP P35520
C	?	-	LYS	DELETION	UNP P35520
C	?	-	SER	DELETION	UNP P35520
C	?	-	SER	DELETION	UNP P35520
C	552	LEU	-	EXPRESSION TAG	UNP P35520
C	553	GLU	-	EXPRESSION TAG	UNP P35520
C	554	HIS	-	EXPRESSION TAG	UNP P35520
C	555	HIS	-	EXPRESSION TAG	UNP P35520
C	556	HIS	-	EXPRESSION TAG	UNP P35520
C	557	HIS	-	EXPRESSION TAG	UNP P35520
C	558	HIS	-	EXPRESSION TAG	UNP P35520
C	559	HIS	-	EXPRESSION TAG	UNP P35520

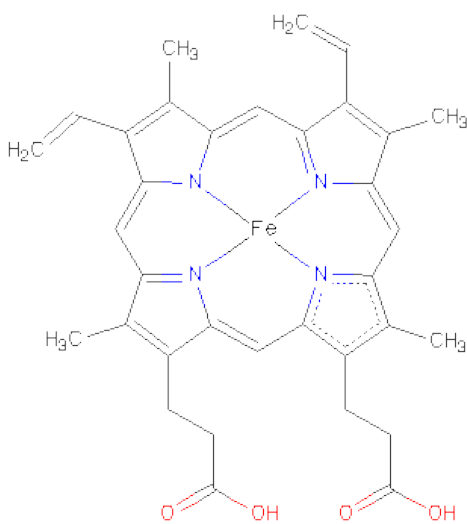
- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: C<sub>8</sub>H<sub>10</sub>NO<sub>6</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	D	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	A	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	C	1	Total	C	N	O	P	0	0
			15	8	1	5	1		

- Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (for-

mula: C<sub>34</sub>H<sub>32</sub>FeN<sub>4</sub>O<sub>4</sub>).



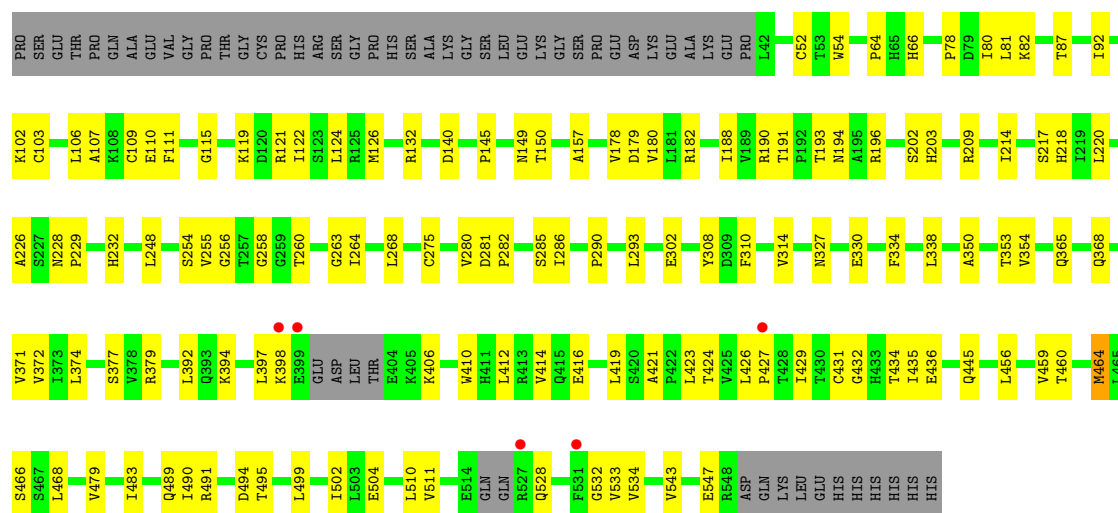
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		



Government	Percentage
Current government	100%
Previous government	0%



Response Category	Percentage
I don't know	1%
Yes	89%
No	8%
Other	2%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	190.45Å 190.45Å 140.85Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	56.62 – 3.39 57.01 – 3.39	Depositor EDS
% Data completeness (in resolution range)	99.7 (56.62-3.39) 99.7 (57.01-3.39)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.85 (at 3.40Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.206 , 0.244 0.208 , 0.244	Depositor DCC
$R_{free}$ test set	2009 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	120.5	Xtriage
Anisotropy	0.195	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 47.4	EDS
Estimated twinning fraction	0.079 for h,-h-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtriage
Outliers	0 of 40187 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15361	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.35% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, PLP

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	0.30	0/3820	0.58	4/5157 (0.1%)
1	B	0.28	0/3861	0.52	1/5225 (0.0%)
1	C	0.29	0/3864	0.55	1/5230 (0.0%)
1	D	0.28	0/3858	0.53	1/5220 (0.0%)
All	All	0.29	0/15403	0.55	7/20832 (0.0%)

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	116	GLY	O-C-N	-8.89	108.48	122.70
1	A	123	SER	O-C-N	-7.32	110.99	122.70
1	A	116	GLY	C-N-CA	6.36	137.61	121.70
1	D	116	GLY	O-C-N	-6.34	112.56	122.70
1	C	124	LEU	O-C-N	-6.30	112.62	122.70

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3757	0	3811	141	1
1	B	3791	0	3845	97	0
1	C	3793	0	3849	112	1
1	D	3788	0	3831	94	0
2	A	15	0	7	4	0
2	B	15	0	7	3	0
2	C	15	0	7	2	0
2	D	15	0	7	3	0
3	A	43	0	30	15	0
3	B	43	0	30	12	0
3	C	43	0	30	14	0
3	D	43	0	30	9	0
All	All	15361	0	15484	443	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 14.

The worst 5 of 443 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:178:VAL:CG2	1:B:190:ARG:HH21	1.36	1.39
1:B:178:VAL:HG21	1:B:190:ARG:NH2	1.29	1.38
1:A:431:CYS:SG	1:A:464:MET:HE1	1.68	1.32
1:B:178:VAL:CG2	1:B:190:ARG:NH2	1.94	1.24
1:A:431:CYS:SG	1:A:464:MET:CE	2.32	1.18

All (1) 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	Distance(Å)	Clash(Å)
1:A:465:LEU:CD2	1:C:421:ALA:CB[5_455]	2.13	0.07

## 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.

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	467/548 (85%)	455 (97%)	12 (3%)	0	100	100
1	B	482/548 (88%)	475 (98%)	7 (2%)	0	100	100
1	C	485/548 (88%)	478 (99%)	7 (1%)	0	100	100
1	D	483/548 (88%)	475 (98%)	7 (1%)	1 (0%)	56	95
All	All	1917/2192 (88%)	1883 (98%)	33 (2%)	1 (0%)	59	96

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	452	ALA

### 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	408/463 (88%)	404 (99%)	4 (1%)	85	96
1	B	413/463 (89%)	407 (98%)	6 (2%)	76	95
1	C	411/463 (89%)	409 (100%)	2 (0%)	94	98
1	D	411/463 (89%)	407 (99%)	4 (1%)	85	96
All	All	1643/1852 (89%)	1627 (99%)	16 (1%)	85	96

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	338	LEU
1	D	414	VAL
1	A	386	LEU
1	D	217	SER
1	A	455	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	194	ASN
1	A	194	ASN
1	A	232	HIS
1	C	194	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

8 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	PLP	A	601	1	14,15,16	1.40	2 (14%)	20,22,23	1.13	1 (5%)
3	HEM	A	602	1	49,50,50	2.25	15 (30%)	46,82,82	2.75	21 (45%)
2	PLP	B	601	1	14,15,16	1.18	1 (7%)	20,22,23	1.12	1 (5%)
3	HEM	B	602	1	49,50,50	2.28	13 (26%)	46,82,82	2.15	9 (19%)
2	PLP	C	601	1	14,15,16	1.24	1 (7%)	20,22,23	1.09	2 (10%)
3	HEM	C	602	1	49,50,50	2.19	16 (32%)	46,82,82	2.48	13 (28%)
2	PLP	D	601	1	14,15,16	1.37	1 (7%)	20,22,23	1.10	1 (5%)
3	HEM	D	602	1	49,50,50	2.29	14 (28%)	46,82,82	1.97	8 (17%)

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	PLP	A	601	1	-	0/6/6/8	0/1/1/1
3	HEM	A	602	1	-	1/14/114/114	0/0/8/8
2	PLP	B	601	1	-	0/6/6/8	0/1/1/1
3	HEM	B	602	1	-	2/14/114/114	0/0/8/8
2	PLP	C	601	1	-	0/6/6/8	0/1/1/1
3	HEM	C	602	1	-	0/14/114/114	0/0/8/8
2	PLP	D	601	1	-	0/6/6/8	0/1/1/1
3	HEM	D	602	1	-	0/14/114/114	0/0/8/8

The worst 5 of 63 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	602	HEM	C3D-C2D	5.90	1.54	1.43
3	D	602	HEM	C3D-C2D	5.88	1.54	1.43
3	B	602	HEM	C3C-C2C	-5.74	1.33	1.43
3	B	602	HEM	C3B-C2B	-5.74	1.33	1.43
3	C	602	HEM	C3C-C2C	-5.66	1.33	1.43

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	602	HEM	C3B-C4B-NB	-10.36	106.59	114.00
3	C	602	HEM	C3B-C4B-NB	-10.22	106.69	114.00
3	B	602	HEM	C3B-C4B-NB	-9.45	107.24	114.00
3	D	602	HEM	C3B-C4B-NB	-7.69	108.50	114.00
3	D	602	HEM	C4D-ND-C1D	6.13	111.44	105.16

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	602	HEM	C2D-C3D-CAD-CBD
3	B	602	HEM	C4D-C3D-CAD-CBD
3	A	602	HEM	C2D-C3D-CAD-CBD

There are no ring outliers.

## 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	485/548 (88%)	0.33	28 (5%) 22 8	43, 83, 160, 205	0
1	B	490/548 (89%)	0.04	6 (1%) 75 39	40, 84, 141, 214	0
1	C	491/548 (89%)	0.07	5 (1%) 79 44	39, 78, 137, 185	0
1	D	491/548 (89%)	0.12	8 (1%) 68 32	34, 75, 159, 207	0
All	All	1957/2192 (89%)	0.14	47 (2%) 56 24	34, 80, 154, 214	0

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	57	GLY	5.9
1	A	443	PHE	4.1
1	A	41	PRO	3.9
1	B	427	PRO	3.8
1	A	456	LEU	3.8

### 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, 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	HEM	C	602	43/43	0.32	2.72	79,79,80,80	0
3	HEM	B	602	43/43	0.28	1.67	112,112,113,113	0
3	HEM	A	602	43/43	0.34	0.97	106,106,107,107	0
3	HEM	D	602	43/43	0.29	0.56	79,79,80,80	0
2	PLP	C	601	15/16	0.26	0.31	45,56,60,63	0
2	PLP	A	601	15/16	0.23	-0.09	50,67,77,77	0
2	PLP	B	601	15/16	0.23	-0.10	57,70,80,81	0
2	PLP	D	601	15/16	0.22	-0.12	49,56,65,66	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.