Full wwPDB X-ray Structure Validation Report (i)

Oct 1, 2024 – 04:22 PM EDT

PDB ID	:	1LH3
Title	:	X-RAY STRUCTURAL INVESTIGATION OF LEGHEMOGLOBIN. VI. STRUCTURE OF ACETATE-FERRILEGHEMOGLOBIN AT A RESOLUTION OF 2.0 ANGSTROMS (RUSSIAN)
Authors	:	Vainshtein, B.K.; Harutyunyan, E.H.; Kuranova, I.P.; Borisov, V.V.; Sosfenov, N.I.; Pavlovsky, A.G.; Grebenko, A.I.; Konareva, N.V.
Deposited on	:	1982-04-23
Resolution	:	2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

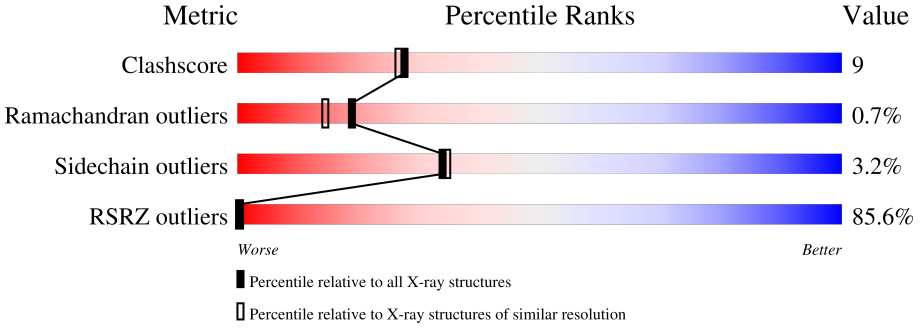
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

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

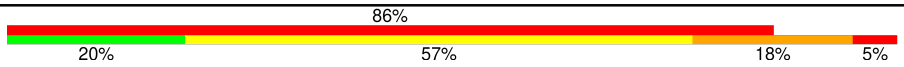
The reported resolution of this entry is 2.00 Å.

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	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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%

Mol	Chain	Length	Quality of chain
1	A	153	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1291 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.

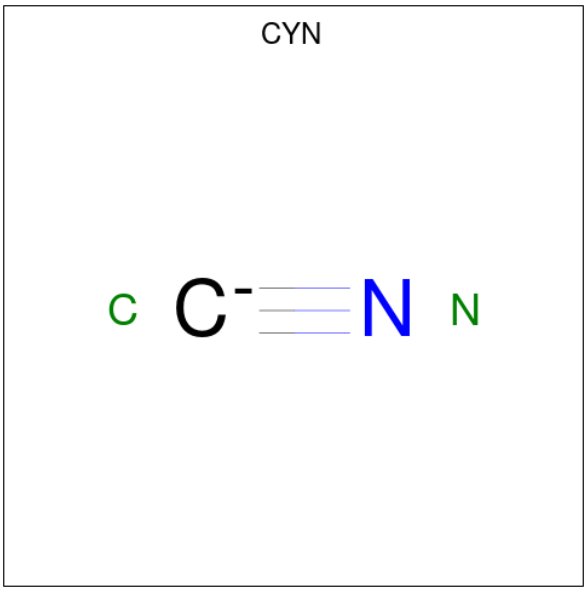
- Molecule 1 is a protein called LEGHEMOGLOBIN (CYANO MET).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	153	1180	761	193	225	1	35	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	79	GLU	GLN	conflict	UNP P02240
A	150	ASP	ASN	conflict	UNP P02240

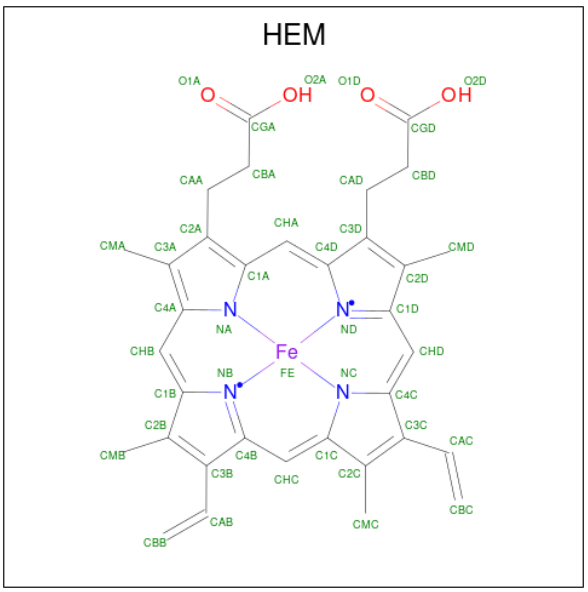
- Molecule 2 is CYANIDE ION (three-letter code: CYN) (formula: CN).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	N		
2	A	1	2	1	1	0	0

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

mula: $C_{34}H_{32}FeN_4O_4$).



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

- Molecule 4 is water.

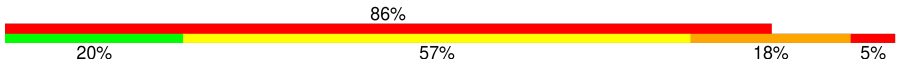
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	66	66	66	0	0

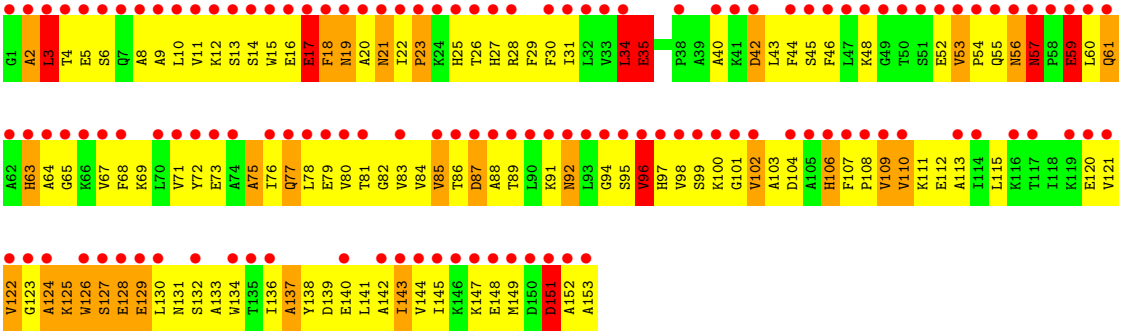
3 Residue-property plots (i)

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. 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. 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: LEGHEMOGLOBIN (CYANO MET)

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	B 1 1 2	Depositor
Cell constants a, b, c, α, β, γ	93.19Å 38.41Å 52.07Å 90.00° 90.00° 98.90°	Depositor
Resolution (Å)	(Not available) – 2.00 46.03 – 2.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-2.00) 92.5 (46.03-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	unknown	Depositor
R, R_{free}	(Not available) , (Not available) 0.440 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	20.4	Xtriage
Anisotropy	0.238	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.44 , 109.6	EDS
L-test for twinning[1]	$< \|L\| > = 0.42$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.65	EDS
Total number of atoms	1291	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.59% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]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: CYN, 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	3.29	152/1214 (12.5%)	2.33	66/1648 (4.0%)

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	10

All (152) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	138	TYR	CZ-OH	11.60	1.57	1.37
1	A	138	TYR	CB-CG	10.13	1.66	1.51
1	A	95	SER	CB-OG	9.96	1.55	1.42
1	A	112	GLU	CG-CD	9.92	1.66	1.51
1	A	15	TRP	CD2-CE2	9.82	1.53	1.41
1	A	120	GLU	CG-CD	9.30	1.66	1.51
1	A	121	VAL	CB-CG2	8.69	1.71	1.52
1	A	65	GLY	C-O	8.67	1.37	1.23
1	A	132[A]	SER	CA-CB	8.66	1.66	1.52
1	A	132[B]	SER	CA-CB	8.66	1.66	1.52
1	A	132[C]	SER	CA-CB	8.66	1.66	1.52
1	A	18	PHE	CB-CG	8.64	1.66	1.51
1	A	94	GLY	CA-C	8.60	1.65	1.51
1	A	106	HIS	CB-CG	8.49	1.65	1.50
1	A	113	ALA	CA-CB	8.26	1.69	1.52
1	A	72	TYR	CG-CD2	8.20	1.49	1.39
1	A	44	PHE	CB-CG	8.16	1.65	1.51
1	A	140	GLU	CB-CG	8.15	1.67	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	123	GLY	CA-C	8.01	1.64	1.51
1	A	144	VAL	CB-CG2	7.88	1.69	1.52
1	A	15	TRP	CB-CG	7.75	1.64	1.50
1	A	72	TYR	CE1-CZ	7.71	1.48	1.38
1	A	11	VAL	CB-CG1	7.61	1.68	1.52
1	A	5	GLU	CD-OE2	7.58	1.33	1.25
1	A	101	GLY	CA-C	7.57	1.64	1.51
1	A	35	GLU	CD-OE2	7.54	1.33	1.25
1	A	14	SER	CB-OG	-7.46	1.32	1.42
1	A	110	VAL	CB-CG2	7.43	1.68	1.52
1	A	88	ALA	N-CA	7.40	1.61	1.46
1	A	79	GLU	CB-CG	7.37	1.66	1.52
1	A	15	TRP	CZ3-CH2	7.33	1.51	1.40
1	A	15	TRP	C-O	7.32	1.37	1.23
1	A	13	SER	CB-OG	-7.32	1.32	1.42
1	A	61	GLN	C-O	7.29	1.37	1.23
1	A	46	PHE	CB-CG	7.15	1.63	1.51
1	A	128	GLU	CD-OE2	-7.09	1.17	1.25
1	A	129	GLU	CD-OE1	7.08	1.33	1.25
1	A	102	VAL	CB-CG1	7.06	1.67	1.52
1	A	124	ALA	C-O	7.03	1.36	1.23
1	A	68	PHE	CE1-CZ	6.99	1.50	1.37
1	A	152	ALA	C-O	6.89	1.36	1.23
1	A	11	VAL	N-CA	6.80	1.59	1.46
1	A	15	TRP	CD1-NE1	6.76	1.49	1.38
1	A	59	GLU	CD-OE2	6.76	1.33	1.25
1	A	68	PHE	CG-CD2	6.74	1.48	1.38
1	A	124	ALA	N-CA	6.68	1.59	1.46
1	A	67	VAL	CB-CG2	6.65	1.66	1.52
1	A	99	SER	N-CA	6.65	1.59	1.46
1	A	85	VAL	CB-CG2	6.63	1.66	1.52
1	A	27	HIS	CE1-NE2	6.60	1.47	1.32
1	A	55	GLN	C-O	6.58	1.35	1.23
1	A	111	LYS	N-CA	6.58	1.59	1.46
1	A	72	TYR	C-O	6.56	1.35	1.23
1	A	13	SER	CA-CB	6.56	1.62	1.52
1	A	73	GLU	CG-CD	6.54	1.61	1.51
1	A	97	HIS	CG-CD2	-6.50	1.24	1.35
1	A	141	LEU	CB-CG	6.47	1.71	1.52
1	A	16	GLU	CG-CD	6.47	1.61	1.51
1	A	27	HIS	CG-ND1	6.46	1.52	1.38
1	A	95	SER	N-CA	6.44	1.59	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	40	ALA	CA-CB	6.43	1.66	1.52
1	A	102	VAL	N-CA	6.40	1.59	1.46
1	A	35	GLU	CD-OE1	-6.39	1.18	1.25
1	A	145	ILE	N-CA	6.39	1.59	1.46
1	A	79	GLU	CD-OE2	6.38	1.32	1.25
1	A	106	HIS	ND1-CE1	6.35	1.50	1.34
1	A	122	VAL	N-CA	6.32	1.58	1.46
1	A	133	ALA	CA-C	6.30	1.69	1.52
1	A	53	VAL	CA-CB	6.30	1.68	1.54
1	A	149	MET	C-O	6.28	1.35	1.23
1	A	77	GLN	CA-CB	6.26	1.67	1.53
1	A	134	TRP	CD2-CE2	-6.26	1.33	1.41
1	A	120	GLU	CA-CB	6.22	1.67	1.53
1	A	30	PHE	N-CA	6.21	1.58	1.46
1	A	52	GLU	N-CA	6.19	1.58	1.46
1	A	71	VAL	CB-CG1	6.18	1.65	1.52
1	A	71	VAL	N-CA	6.18	1.58	1.46
1	A	15	TRP	CG-CD1	6.15	1.45	1.36
1	A	45	SER	N-CA	6.09	1.58	1.46
1	A	8	ALA	C-O	6.08	1.34	1.23
1	A	63	HIS	CA-C	6.07	1.68	1.52
1	A	68	PHE	C-O	6.05	1.34	1.23
1	A	86	THR	N-CA	6.02	1.58	1.46
1	A	84	VAL	N-CA	6.02	1.58	1.46
1	A	18	PHE	N-CA	6.01	1.58	1.46
1	A	148	GLU	CB-CG	5.98	1.63	1.52
1	A	43	LEU	CA-CB	5.94	1.67	1.53
1	A	2	ALA	CA-CB	5.92	1.64	1.52
1	A	130	LEU	N-CA	5.87	1.58	1.46
1	A	97	HIS	CA-CB	5.85	1.66	1.53
1	A	82	GLY	CA-C	5.83	1.61	1.51
1	A	76	ILE	C-O	5.83	1.34	1.23
1	A	81	THR	N-CA	-5.82	1.34	1.46
1	A	137	ALA	CA-C	5.81	1.68	1.52
1	A	44	PHE	CD2-CE2	5.80	1.50	1.39
1	A	115	LEU	N-CA	5.78	1.57	1.46
1	A	99	SER	C-O	5.78	1.34	1.23
1	A	127	SER	CA-CB	5.77	1.61	1.52
1	A	148	GLU	C-O	5.75	1.34	1.23
1	A	20	ALA	N-CA	-5.74	1.34	1.46
1	A	152	ALA	N-CA	5.72	1.57	1.46
1	A	125	LYS	CD-CE	5.63	1.65	1.51

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	148	GLU	N-CA	5.60	1.57	1.46
1	A	52	GLU	CG-CD	-5.59	1.43	1.51
1	A	44	PHE	CD1-CE1	5.59	1.50	1.39
1	A	108	PRO	C-N	5.58	1.46	1.34
1	A	45	SER	CB-OG	5.55	1.49	1.42
1	A	75	ALA	C-O	5.53	1.33	1.23
1	A	148	GLU	CD-OE1	5.53	1.31	1.25
1	A	104	ASP	N-CA	5.52	1.57	1.46
1	A	148	GLU	CD-OE2	5.51	1.31	1.25
1	A	83	VAL	N-CA	5.50	1.57	1.46
1	A	46	PHE	N-CA	-5.50	1.35	1.46
1	A	6	SER	CA-CB	5.49	1.61	1.52
1	A	75	ALA	N-CA	5.45	1.57	1.46
1	A	92	ASN	CB-CG	5.45	1.63	1.51
1	A	110	VAL	CA-C	5.44	1.67	1.52
1	A	4	THR	C-O	5.43	1.33	1.23
1	A	147	LYS	CA-CB	5.42	1.65	1.53
1	A	143	ILE	CA-CB	5.41	1.67	1.54
1	A	16	GLU	CD-OE1	5.39	1.31	1.25
1	A	68	PHE	CG-CD1	5.37	1.46	1.38
1	A	96	VAL	C-O	5.37	1.33	1.23
1	A	68	PHE	CA-CB	-5.37	1.42	1.53
1	A	128	GLU	CG-CD	5.36	1.59	1.51
1	A	127	SER	C-O	5.35	1.33	1.23
1	A	98	VAL	CB-CG2	5.34	1.64	1.52
1	A	111	LYS	CB-CG	5.33	1.67	1.52
1	A	20	ALA	CA-CB	5.31	1.63	1.52
1	A	138	TYR	CE2-CZ	-5.31	1.31	1.38
1	A	85	VAL	CA-C	5.26	1.66	1.52
1	A	18	PHE	CD2-CE2	5.25	1.49	1.39
1	A	69	LYS	C-O	5.25	1.33	1.23
1	A	109	VAL	CA-CB	5.24	1.65	1.54
1	A	100	LYS	C-N	-5.21	1.23	1.33
1	A	107	PHE	N-CA	5.21	1.56	1.46
1	A	77	GLN	N-CA	-5.20	1.35	1.46
1	A	68	PHE	N-CA	5.18	1.56	1.46
1	A	136	ILE	N-CA	-5.18	1.35	1.46
1	A	30	PHE	CG-CD2	-5.18	1.30	1.38
1	A	18	PHE	CD1-CE1	5.14	1.49	1.39
1	A	81	THR	CA-CB	5.13	1.66	1.53
1	A	45	SER	C-O	5.13	1.33	1.23
1	A	98	VAL	CA-C	5.10	1.66	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	153	ALA	C-OXT	5.09	1.33	1.23
1	A	140	GLU	CD-OE2	5.07	1.31	1.25
1	A	29	PHE	CG-CD2	5.05	1.46	1.38
1	A	126	TRP	CZ2-CH2	5.05	1.47	1.37
1	A	79	GLU	C-O	5.03	1.32	1.23
1	A	87	ASP	CA-C	5.03	1.66	1.52
1	A	9	ALA	N-CA	-5.01	1.36	1.46
1	A	96	VAL	CB-CG1	5.01	1.63	1.52

All (66) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	72	TYR	CB-CG-CD1	-9.49	115.31	121.00
1	A	112	GLU	OE1-CD-OE2	-8.73	112.82	123.30
1	A	15	TRP	CG-CD2-CE3	-8.68	126.09	133.90
1	A	80	VAL	CA-CB-CG2	-8.56	98.07	110.90
1	A	16	GLU	OE1-CD-OE2	-8.54	113.05	123.30
1	A	124	ALA	CB-CA-C	-8.25	97.73	110.10
1	A	134	TRP	CG-CD1-NE1	-7.83	102.27	110.10
1	A	29	PHE	CB-CG-CD1	-7.72	115.39	120.80
1	A	73	GLU	OE1-CD-OE2	-7.57	114.21	123.30
1	A	126	TRP	CE2-CD2-CG	-7.35	101.42	107.30
1	A	17	GLU	OE1-CD-OE2	-7.27	114.58	123.30
1	A	54	PRO	N-CA-CB	7.19	111.93	103.30
1	A	128	GLU	OE1-CD-OE2	-7.15	114.72	123.30
1	A	110	VAL	O-C-N	-6.77	111.86	122.70
1	A	34	LEU	CB-CG-CD2	6.77	122.51	111.00
1	A	30	PHE	CB-CG-CD2	-6.77	116.06	120.80
1	A	28	ARG	NE-CZ-NH2	-6.67	116.97	120.30
1	A	46	PHE	CB-CG-CD2	6.67	125.47	120.80
1	A	72	TYR	CG-CD1-CE1	-6.63	116.00	121.30
1	A	120	GLU	OE1-CD-OE2	-6.56	115.43	123.30
1	A	29	PHE	CD1-CG-CD2	6.50	126.76	118.30
1	A	48	LYS	CB-CA-C	-6.47	97.47	110.40
1	A	15	TRP	CD2-CE3-CZ3	-6.36	110.53	118.80
1	A	15	TRP	CH2-CZ2-CE2	-6.33	111.07	117.40
1	A	68	PHE	CB-CG-CD1	-6.29	116.40	120.80
1	A	134	TRP	CD1-CG-CD2	6.24	111.29	106.30
1	A	15	TRP	CD1-NE1-CE2	-6.21	103.41	109.00
1	A	138	TYR	CB-CG-CD2	6.11	124.67	121.00
1	A	124	ALA	N-CA-CB	6.06	118.58	110.10
1	A	53	VAL	CA-CB-CG2	-6.05	101.83	110.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	78	LEU	CB-CG-CD2	-5.96	100.87	111.00
1	A	133	ALA	O-C-N	-5.96	113.17	122.70
1	A	15	TRP	CE2-CD2-CE3	5.95	125.84	118.70
1	A	127	SER	N-CA-CB	5.94	119.41	110.50
1	A	131	ASN	O-C-N	5.93	132.18	122.70
1	A	45	SER	O-C-N	5.91	132.15	122.70
1	A	76	ILE	O-C-N	5.90	132.14	122.70
1	A	126	TRP	NE1-CE2-CD2	5.77	113.07	107.30
1	A	61	GLN	O-C-N	5.69	131.81	122.70
1	A	134	TRP	CD1-NE1-CE2	5.68	114.12	109.00
1	A	23	PRO	N-CA-CB	5.63	110.05	103.30
1	A	15	TRP	CB-CG-CD1	5.62	134.31	127.00
1	A	12	LYS	O-C-N	5.62	131.69	122.70
1	A	53	VAL	CA-CB-CG1	5.62	119.32	110.90
1	A	138	TYR	CD1-CE1-CZ	5.59	124.83	119.80
1	A	5	GLU	OE1-CD-OE2	-5.56	116.63	123.30
1	A	104	ASP	CB-CA-C	-5.54	99.31	110.40
1	A	42	ASP	CB-CG-OD1	-5.49	113.36	118.30
1	A	126	TRP	CD1-CG-CD2	5.45	110.66	106.30
1	A	83	VAL	CA-CB-CG1	-5.44	102.74	110.90
1	A	142	ALA	O-C-N	5.38	131.32	122.70
1	A	122	VAL	CA-CB-CG1	-5.30	102.95	110.90
1	A	151	ASP	CB-CG-OD2	-5.22	113.60	118.30
1	A	87	ASP	CB-CG-OD2	-5.22	113.60	118.30
1	A	126	TRP	CE2-CD2-CE3	5.17	124.90	118.70
1	A	8	ALA	O-C-N	5.17	130.97	122.70
1	A	106	HIS	CG-ND1-CE1	5.14	115.40	108.20
1	A	72	TYR	CZ-CE2-CD2	-5.14	115.17	119.80
1	A	89	THR	N-CA-CB	-5.08	100.64	110.30
1	A	15	TRP	NE1-CE2-CZ2	-5.08	124.81	130.40
1	A	77	GLN	O-C-N	-5.07	114.58	122.70
1	A	57	ASN	CB-CG-OD1	-5.07	111.47	121.60
1	A	3	LEU	CB-CA-C	5.06	119.81	110.20
1	A	19	ASN	N-CA-CB	-5.04	101.53	110.60
1	A	29	PHE	CG-CD1-CE1	-5.02	115.28	120.80
1	A	60	LEU	CB-CG-CD2	5.01	119.52	111.00

There are no chirality outliers.

All (10) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	110	VAL	Mainchain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	151	ASP	Sidechain
1	A	17	GLU	Sidechain
1	A	19	ASN	Sidechain
1	A	35	GLU	Sidechain
1	A	42	ASP	Sidechain
1	A	56	ASN	Sidechain
1	A	57	ASN	Sidechain
1	A	59	GLU	Sidechain
1	A	61	GLN	Sidechain

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	1180	0	1200	21	18
2	A	2	0	0	0	0
3	A	43	0	30	2	0
4	A	66	0	0	0	4
All	All	1291	0	1230	21	20

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 9.

All (21) 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:77:GLN:NE2	1:A:85:VAL:H	1.98	0.61
1:A:31:ILE:O	1:A:35:GLU:HG3	2.01	0.60
1:A:92:ASN:O	1:A:96:VAL:HG12	2.05	0.57
1:A:87:ASP:O	1:A:91:LYS:HG3	2.04	0.57
1:A:106:HIS:O	1:A:109:VAL:HB	2.06	0.56
1:A:21:ASN:C	1:A:21:ASN:HD22	2.12	0.53
1:A:26:THR:HB	1:A:64:ALA:HB3	1.90	0.52
1:A:18:PHE:CE1	1:A:25:HIS:HB3	2.45	0.51
1:A:34:LEU:HD22	1:A:53:VAL:HG21	1.92	0.50
1:A:21:ASN:C	1:A:21:ASN:ND2	2.67	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:17:GLU:OE2	1:A:122:VAL:HG12	2.15	0.47
1:A:75:ALA:HB2	1:A:137:ALA:HB2	1.97	0.47
1:A:2:ALA:O	1:A:3:LEU:HB2	2.14	0.47
1:A:22:ILE:HB	1:A:23:PRO:HD3	1.97	0.46
1:A:63:HIS:HE1	3:A:155:HEM:CHA	2.29	0.45
1:A:139:ASP:O	1:A:143:ILE:HG13	2.16	0.45
1:A:103:ALA:O	1:A:106:HIS:HB2	2.20	0.42
1:A:57:ASN:OD1	1:A:59:GLU:HB2	2.20	0.42
1:A:102:VAL:HG13	3:A:155:HEM:HAC	2.01	0.41
1:A:34:LEU:HD22	1:A:53:VAL:CG2	2.51	0.41
1:A:75:ALA:HB2	1:A:137:ALA:CB	2.51	0.41

All (20) 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:A:126:TRP:O	1:A:126:TRP:O[2_555]	1.03	1.17
1:A:124:ALA:CB	1:A:129:GLU:N[2_555]	1.20	1.00
1:A:124:ALA:O	1:A:127:SER:CB[2_555]	1.23	0.97
4:A:218:HOH:O	4:A:218:HOH:O[2_555]	1.34	0.86
1:A:124:ALA:O	1:A:127:SER:CA[2_555]	1.36	0.84
1:A:124:ALA:C	1:A:127:SER:CB[2_555]	1.42	0.78
1:A:124:ALA:O	1:A:127:SER:C[2_555]	1.46	0.74
1:A:124:ALA:O	1:A:128:GLU:N[2_555]	1.64	0.56
1:A:125:LYS:CA	1:A:127:SER:OG[2_555]	1.68	0.52
1:A:125:LYS:CA	1:A:127:SER:CB[2_555]	1.79	0.41
1:A:125:LYS:N	1:A:127:SER:CB[2_555]	1.82	0.38
1:A:10:LEU:CD1	1:A:125:LYS:CD[2_555]	1.83	0.37
1:A:56:ASN:OD1	4:A:171:HOH:O[1_565]	1.86	0.34
1:A:124:ALA:O	1:A:127:SER:OG[2_555]	1.86	0.34
1:A:124:ALA:CB	1:A:128:GLU:C[2_555]	2.02	0.18
1:A:125:LYS:C	1:A:127:SER:CB[2_555]	2.02	0.18
4:A:192:HOH:O	4:A:218:HOH:O[2_555]	2.07	0.13
1:A:56:ASN:N	4:A:171:HOH:O[1_565]	2.13	0.07
1:A:124:ALA:C	1:A:127:SER:OG[2_555]	2.17	0.03
1:A:125:LYS:N	1:A:127:SER:OG[2_555]	2.17	0.03

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	153/153 (100%)	148 (97%)	4 (3%)	1 (1%)	19 14

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	LEU

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	127/125 (102%)	123 (97%)	4 (3%)	35 36

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	21	ASN
1	A	34	LEU
1	A	96	VAL
1	A	151	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	21	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	25	HIS
1	A	61	GLN
1	A	77	GLN
1	A	92	ASN

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)

2 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	$\#\|Z\| > 2$
3	HEM	A	155	2,1	42,50,50	4.10	31 (73%)	46,82,82	2.60	20 (43%)
2	CYN	A	154	3	1,1,1	0.20	0	-		

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
3	HEM	A	155	2,1	-	4/12/54/54	-

All (31) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	155	HEM	C1D-C2D	8.00	1.60	1.44
3	A	155	HEM	C1A-NA	7.82	1.52	1.36
3	A	155	HEM	C4D-C3D	7.31	1.57	1.45
3	A	155	HEM	C3B-C4B	7.06	1.58	1.44
3	A	155	HEM	CAB-C3B	6.55	1.64	1.47
3	A	155	HEM	CBD-CGD	6.54	1.65	1.50
3	A	155	HEM	FE-NB	6.21	2.32	1.98
3	A	155	HEM	C1B-C2B	5.83	1.56	1.44
3	A	155	HEM	CMD-C2D	5.39	1.61	1.50
3	A	155	HEM	FE-ND	4.71	2.24	1.98
3	A	155	HEM	C4A-CHB	4.58	1.53	1.41
3	A	155	HEM	CAD-C3D	4.56	1.63	1.51
3	A	155	HEM	C3C-C4C	4.47	1.48	1.41
3	A	155	HEM	CMB-C2B	4.36	1.59	1.50
3	A	155	HEM	C3C-C2C	4.32	1.46	1.40
3	A	155	HEM	C2A-C3A	4.17	1.50	1.37
3	A	155	HEM	C4A-NA	3.99	1.44	1.36
3	A	155	HEM	O1A-CGA	3.85	1.34	1.22
3	A	155	HEM	C3C-CAC	3.76	1.56	1.47
3	A	155	HEM	CMC-C2C	3.73	1.60	1.51
3	A	155	HEM	C1A-CHA	3.26	1.50	1.41
3	A	155	HEM	CBA-CGA	3.24	1.58	1.50
3	A	155	HEM	CAA-C2A	2.91	1.59	1.52
3	A	155	HEM	CMA-C3A	2.90	1.57	1.51
3	A	155	HEM	CHB-C1B	2.72	1.41	1.34
3	A	155	HEM	O1D-CGD	2.58	1.30	1.22
3	A	155	HEM	CHA-C4D	2.52	1.40	1.34
3	A	155	HEM	C4D-ND	2.50	1.44	1.40
3	A	155	HEM	CBD-CAD	-2.43	1.43	1.51
3	A	155	HEM	CBB-CAB	2.34	1.41	1.30
3	A	155	HEM	C4B-NB	2.16	1.43	1.38

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	155	HEM	CMA-C3A-C4A	-6.78	118.52	128.46
3	A	155	HEM	C4A-C3A-C2A	6.25	111.35	107.00
3	A	155	HEM	C2C-C3C-C4C	5.81	110.95	106.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	155	HEM	C4C-CHD-C1D	5.25	129.49	122.56
3	A	155	HEM	CAD-CBD-CGD	-4.13	102.72	113.67
3	A	155	HEM	C3C-C4C-NC	-3.85	103.67	110.94
3	A	155	HEM	CAD-C3D-C4D	3.34	130.51	124.70
3	A	155	HEM	C3B-C4B-NB	3.33	111.86	109.47
3	A	155	HEM	O2D-CGD-O1D	-3.08	115.41	123.33
3	A	155	HEM	C4B-C3B-C2B	-2.97	104.55	107.28
3	A	155	HEM	CHD-C1D-ND	-2.87	121.35	124.44
3	A	155	HEM	CMA-C3A-C2A	2.74	130.11	124.94
3	A	155	HEM	CBA-CAA-C2A	-2.74	107.93	112.54
3	A	155	HEM	CHB-C1B-NB	-2.73	120.99	124.37
3	A	155	HEM	CMD-C2D-C1D	2.67	129.20	125.03
3	A	155	HEM	CMC-C2C-C3C	2.51	129.70	124.68
3	A	155	HEM	CHA-C4D-ND	-2.35	121.46	124.37
3	A	155	HEM	C4B-CHC-C1C	2.32	125.62	122.56
3	A	155	HEM	CAD-C3D-C2D	-2.17	123.80	127.87
3	A	155	HEM	O2D-CGD-CBD	2.06	120.50	114.00

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	155	HEM	CAA-CBA-CGA-O1A
3	A	155	HEM	CAA-CBA-CGA-O2A
3	A	155	HEM	CAD-CBD-CGD-O1D
3	A	155	HEM	CAD-CBD-CGD-O2D

There are no ring outliers.

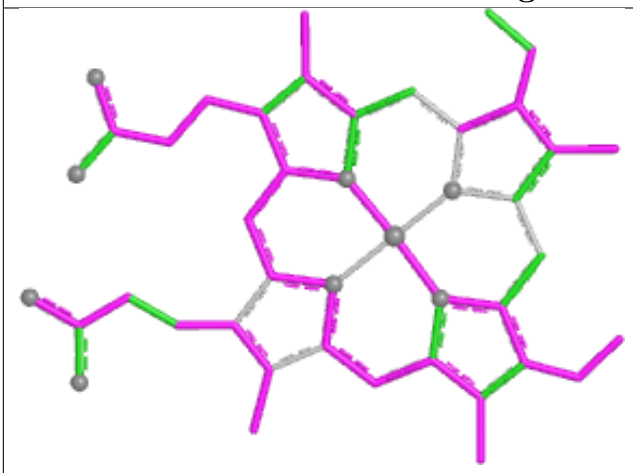
1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	155	HEM	2	0

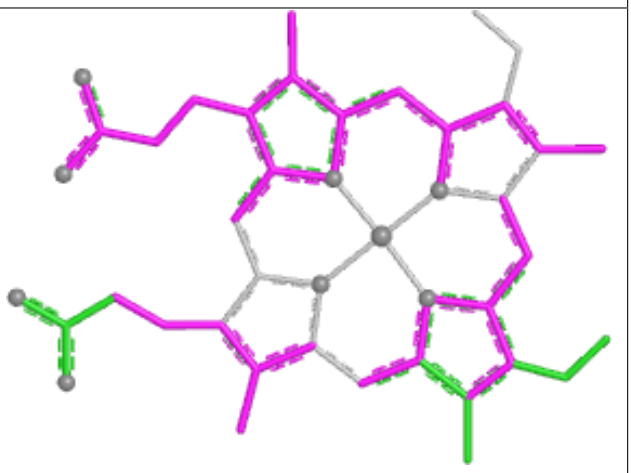
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

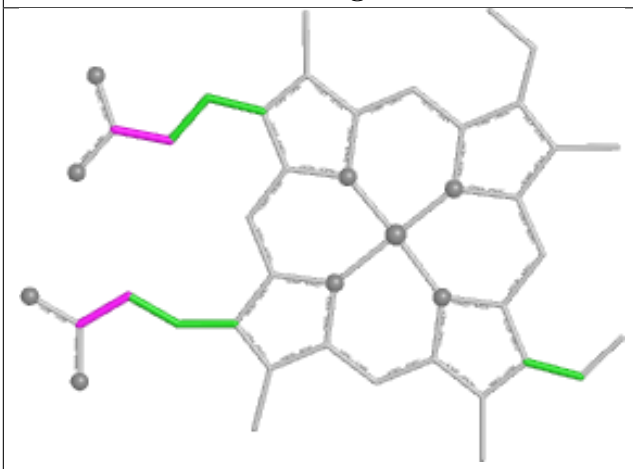
Ligand HEM A 155



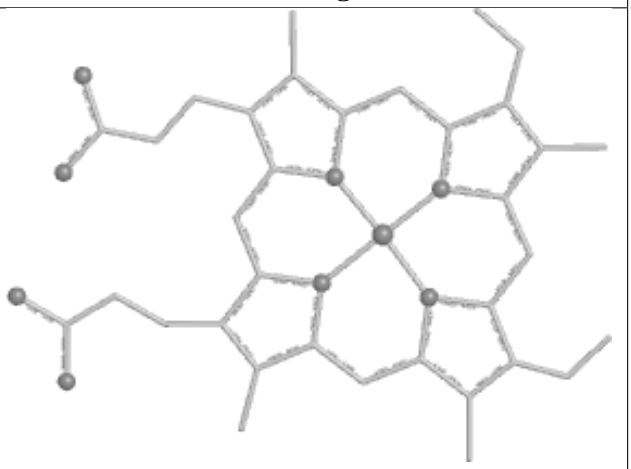
Bond lengths



Bond angles



Torsions



Rings

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)

Warning: The R factor obtained from EDS is 0.4431,which does not match the depositor's R factor of 0.0. Please interpret the results in this section carefully.

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	153/153 (100%)	3.60	131 (85%)	0	0	7, 17, 39, 54	19 (12%)

All (131) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	GLY	13.4
1	A	152	ALA	8.6
1	A	48	LYS	8.1
1	A	9	ALA	7.9
1	A	2	ALA	7.6
1	A	153	ALA	7.2
1	A	45	SER	6.8
1	A	88	ALA	6.8
1	A	49	GLY	6.8
1	A	126	TRP	6.7
1	A	83	VAL	6.5
1	A	123	GLY	6.5
1	A	50	THR	6.1
1	A	150	ASP	6.1
1	A	136	ILE	6.0
1	A	62	ALA	5.9
1	A	128	GLU	5.7
1	A	4	THR	5.6
1	A	58	PRO	5.6
1	A	92	ASN	5.6
1	A	147	LYS	5.5
1	A	151	ASP	5.5
1	A	56	ASN	5.5
1	A	142	ALA	5.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	149	MET	5.5
1	A	86	THR	5.4
1	A	3	LEU	5.4
1	A	90	LEU	5.3
1	A	23	PRO	5.2
1	A	143	ILE	5.2
1	A	5	GLU	5.1
1	A	79	GLU	5.1
1	A	20	ALA	4.8
1	A	53	VAL	4.8
1	A	129	GLU	4.6
1	A	54	PRO	4.5
1	A	130	LEU	4.4
1	A	26	THR	4.4
1	A	6	SER	4.4
1	A	81	THR	4.3
1	A	107	PHE	4.3
1	A	96	VAL	4.2
1	A	91	LYS	4.1
1	A	15	TRP	4.1
1	A	10	LEU	4.1
1	A	51	SER	4.1
1	A	78	LEU	4.0
1	A	59	GLU	3.9
1	A	85	VAL	3.9
1	A	144	VAL	3.9
1	A	148	GLU	3.9
1	A	27	HIS	3.8
1	A	64	ALA	3.8
1	A	146	LYS	3.8
1	A	21	ASN	3.8
1	A	74	ALA	3.8
1	A	14	SER	3.8
1	A	55	GLN	3.7
1	A	18	PHE	3.7
1	A	57	ASN	3.7
1	A	67	VAL	3.6
1	A	95	SER	3.6
1	A	94	GLY	3.6
1	A	132[A]	SER	3.6
1	A	22	ILE	3.6
1	A	12	LYS	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	44	PHE	3.6
1	A	47	LEU	3.5
1	A	61	GLN	3.5
1	A	114	ILE	3.5
1	A	97	HIS	3.5
1	A	108	PRO	3.4
1	A	25	HIS	3.4
1	A	71	VAL	3.4
1	A	119	LYS	3.4
1	A	11	VAL	3.3
1	A	89	THR	3.3
1	A	42	ASP	3.3
1	A	98	VAL	3.3
1	A	99	SER	3.2
1	A	127	SER	3.2
1	A	117	THR	3.2
1	A	120	GLU	3.2
1	A	31	ILE	3.2
1	A	52	GLU	3.1
1	A	33	VAL	3.1
1	A	87	ASP	3.0
1	A	46	PHE	2.9
1	A	124	ALA	2.9
1	A	102	VAL	2.9
1	A	19	ASN	2.9
1	A	122	VAL	2.8
1	A	77	GLN	2.8
1	A	60	LEU	2.8
1	A	104	ASP	2.8
1	A	105	ALA	2.8
1	A	93	LEU	2.8
1	A	65	GLY	2.7
1	A	16	GLU	2.7
1	A	100	LYS	2.7
1	A	121	VAL	2.7
1	A	24	LYS	2.6
1	A	70	LEU	2.6
1	A	80	VAL	2.6
1	A	145	ILE	2.6
1	A	140	GLU	2.6
1	A	73	GLU	2.5
1	A	72	TYR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	34	LEU	2.5
1	A	8	ALA	2.5
1	A	110	VAL	2.5
1	A	66	LYS	2.4
1	A	41	LYS	2.4
1	A	68	PHE	2.4
1	A	17	GLU	2.4
1	A	101	GLY	2.3
1	A	113	ALA	2.3
1	A	116	LYS	2.3
1	A	76	ILE	2.2
1	A	40	ALA	2.2
1	A	135	THR	2.2
1	A	13	SER	2.2
1	A	28	ARG	2.2
1	A	30	PHE	2.1
1	A	32	LEU	2.1
1	A	109	VAL	2.1
1	A	38	PRO	2.1
1	A	63	HIS	2.0
1	A	7	GLN	2.0
1	A	134	TRP	2.0
1	A	106	HIS	2.0

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 monosaccharides 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.

Continued on next page...

Continued from previous page...

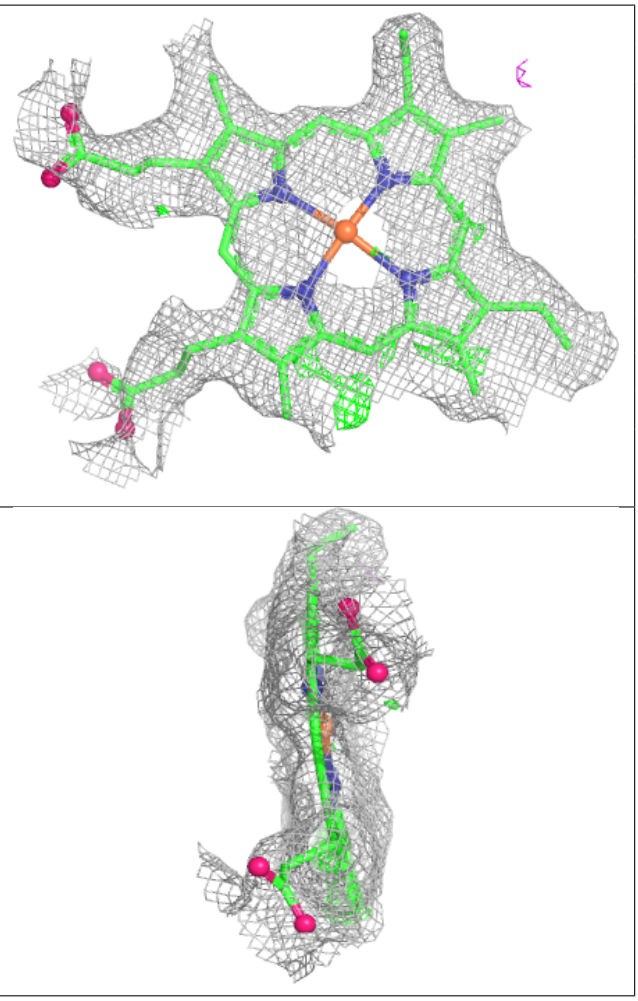
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	HEM	A	155	43/43	0.85	0.18	5,14,50,61	0
2	CYN	A	154	2/2	0.88	0.24	3,3,3,15	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 HEM A 155:

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.