

# Full wwPDB X-ray Structure Validation Report (i)

#### Feb 3, 2024 – 06:47 PM EST

PDB ID	:	1MOD
Title	:	HIGH-RESOLUTION CRYSTAL STRUCTURES OF DISTAL HISTIDINE
		MUTANTS OF SPERM WHALE MYOGLOBIN
Authors	:	Quillin, M.L.; Phillips Jr., G.N.
Deposited on	:	1994-12-01
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 (1)) were used in the production of this report:

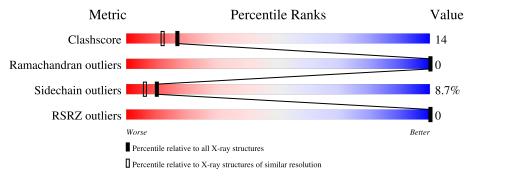
MolProbity Mogul Xtriage (Phenix) EDS	:	4.02b-467 1.8.5 (274361), CSD as541be (2020) 1.13 2.36
buster-report Percentile statistics Refmac	: : :	1.1.7 (2018) 20191225.v01 (using entries in the PDB archive December 25th 2019) 5.8.0158 7.0.044 (Gargrove)
Ideal geometry (DNA, RNA) Validation Pipeline (wwPDB-VP)		Parkinson et al. (1996) 2.36

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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% 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 a	chain	
1	А	154	57%	31%	9% •



#### 1MOD

# 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1398 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.

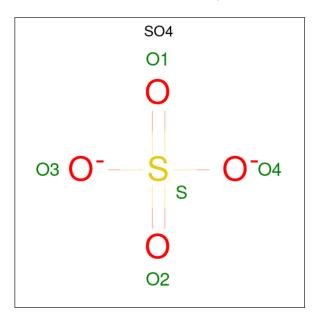
• Molecule 1 is a protein called MYOGLOBIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	154	Total 1222	C 786	N 216	0 217	${ m S} { m 3}$	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	64	THR	HIS	conflict	UNP P02185
А	122	ASN	ASP	conflict	UNP P02185

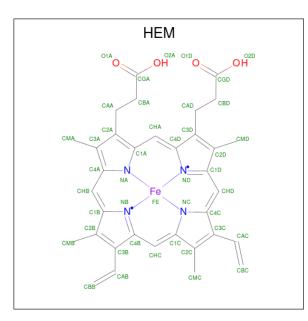
• Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Ato	oms		ZeroOcc	AltConf
2	А	1	Total 5	0 4	S 1	0	0

• Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).





Mol	Chain	Residues		Ate	oms			ZeroOcc	AltConf
2	Λ	1	Total	С	Fe	Ν	0	0	0
5	Л	1	43	34	1	4	4	0	0

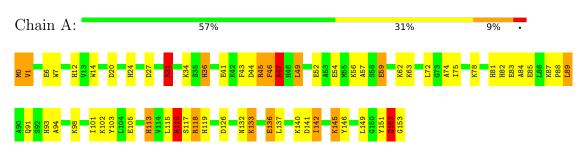
• Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	128	Total         O           128         128	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: MYOGLOBIN



# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 6	Depositor
Cell constants	91.20Å 91.20Å 45.87Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $120.00^{\circ}$	Depositor
Resolution (Å)	(Not available) - 2.00	Depositor
Resolution (A)	5.00 - 2.00	EDS
% Data completeness	(Not available) ((Not available)- $2.00$ )	Depositor
(in resolution range)	73.8(5.00-2.00)	EDS
R <sub>merge</sub>	(Not available)	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	X-PLOR	Depositor
$R, R_{free}$	0.141 , (Not available)	Depositor
$\mathbf{n}, \mathbf{n}_{free}$	0.141 , (Not available)	DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor $(Å^2)$	18.3	Xtriage
Anisotropy	0.120	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.52 , $121.4$	EDS
L-test for twinning <sup>1</sup>	$<  L  > = 0.47, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.038 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	1398	wwPDB-VP
Average B, all atoms $(Å^2)$	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows:

<sup>&</sup>lt;sup>1</sup>Theoretical 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 (i)

## 5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, 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	Bo	nd lengths	Bond angles		
	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.95	5/1249~(0.4%)	1.36	13/1676~(0.8%)	

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	А	0	32

Mol	Chain	Res	Type	Atoms	Z	Observed(A)	Ideal(Å)
1	А	113	HIS	CB-CG	5.66	1.60	1.50
1	А	119	HIS	CG-CD2	5.56	1.45	1.35
1	А	36	HIS	CB-CG	5.17	1.59	1.50
1	А	24	HIS	CG-CD2	5.13	1.44	1.35
1	А	116	HIS	CB-CG	5.01	1.59	1.50

All (5) bond length outliers are listed below:

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$Observed(^{o})$	$Ideal(^{o})$
1	А	7	TRP	CD1-CG-CD2	8.37	113.00	106.30
1	А	118	ARG	NE-CZ-NH2	-8.04	116.28	120.30
1	А	14	TRP	CD1-CG-CD2	7.10	111.98	106.30
1	А	31	ARG	NE-CZ-NH2	-6.95	116.83	120.30
1	А	7	TRP	CE2-CD2-CG	-6.94	101.75	107.30
1	А	116	HIS	CG-CD2-NE2	-6.76	96.36	109.20
1	А	14	TRP	CE2-CD2-CG	-6.59	102.03	107.30
1	А	24	HIS	CB-CG-ND1	6.58	139.64	123.20

Continued on next page...



Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	А	116	HIS	ND1-CG-CD2	5.93	117.10	108.80
1	А	12	HIS	CB-CG-ND1	5.78	137.64	123.20
1	А	113	HIS	CA-CB-CG	5.60	123.13	113.60
1	А	113	HIS	CG-CD2-NE2	-5.33	99.08	109.20
1	А	93	HIS	CA-CB-CG	-5.05	105.02	113.60

Continued from previous page...

There are no chirality outliers.

All $(32)$ planarity outliers are listed below:	
---	--

Mol	Chain	Res	Type	Group
1	А	0	MET	Mainchain
1	А	1	VAL	Mainchain
1	А	103	TYR	Mainchain
1	А	105	GLU	Sidechain
1	А	113	HIS	Sidechain
1	А	116	HIS	Mainchain
1	А	118	ARG	Mainchain
1	А	126	ASP	Sidechain
1	А	136	GLU	Sidechain
1	А	146	TYR	Sidechain
1	А	152	GLN	Mainchain
1	А	20	ASP	Sidechain
1	А	27	ASP	Mainchain
1	А	31	ARG	Sidechain
1	А	34	LYS	Mainchain
1	А	36	HIS	Mainchain
1	А	43	PHE	Mainchain
1	А	44	ASP	Sidechain
1	А	46	PHE	Mainchain
1	А	47	LYS	Mainchain
1	А	49	LEU	Mainchain
1	А	54	GLU	Sidechain
1	А	57	ALA	Mainchain
1	А	72	LEU	Mainchain
1	А	74	ALA	Mainchain
1	А	75	ILE	Mainchain
1	А	78	LYS	Mainchain
1	А	82	HIS	Mainchain
1	А	84	ALA	Mainchain
1	А	85	GLU	Sidechain
1	А	89	LEU	Mainchain
1	А	94	ALA	Mainchain



### 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	А	1222	0	1253	36	0
2	А	5	0	0	0	0
3	А	43	0	30	0	0
4	А	128	0	0	11	1
All	All	1398	0	1283	36	1

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 14.

All (36) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

A + 1	A + 0	Interatomic	Clash
Atom-1	Atom-2	distance $(\text{\AA})$	overlap (Å)
1:A:63:LYS:HD3	4:A:311:HOH:O	1.51	1.09
1:A:145:LYS:N	1:A:145:LYS:HD3	1.84	0.88
1:A:141:ASP:O	1:A:145:LYS:HE2	1.75	0.86
1:A:98:LYS:HE3	4:A:302:HOH:O	1.81	0.81
1:A:101:ILE:HG12	1:A:153:GLY:OXT	1.84	0.76
1:A:132:ASN:O	1:A:136:GLU:HG3	1.86	0.76
1:A:87:LYS:HB2	1:A:88:PRO:HD3	1.73	0.71
1:A:52:GLU:OE2	1:A:56:LYS:HE3	1.93	0.69
1:A:83:GLU:OE1	4:A:280:HOH:O	2.10	0.68
1:A:152:GLN:OE1	1:A:152:GLN:HA	1.95	0.65
1:A:31:ARG:HD2	4:A:260:HOH:O	1.99	0.62
1:A:59:GLU:O	1:A:63:LYS:HG3	2.01	0.61
1:A:0:MET:C	1:A:1:VAL:HG23	2.24	0.58
1:A:141:ASP:O	1:A:145:LYS:CE	2.50	0.56
1:A:0:MET:HG3	1:A:137:LEU:HD13	1.88	0.56
1:A:45:ARG:HB2	4:A:248:HOH:O	2.07	0.54
1:A:6:GLU:OE1	1:A:133:LYS:HE2	2.08	0.54
1:A:141:ASP:C	1:A:145:LYS:HE2	2.30	0.52
1:A:31:ARG:C	1:A:31:ARG:HD3	2.31	0.51
1:A:87:LYS:HB2	1:A:88:PRO:CD	2.40	0.50
1:A:151:TYR:O	1:A:152:GLN:C	2.48	0.50
1:A:151:TYR:O	1:A:152:GLN:O	2.30	0.50
1:A:46:PHE:HB2	1:A:49:LEU:HD11	1.94	0.50

Continued on next page...



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:116:HIS:HD2	4:A:258:HOH:O	1.94	0.50
1:A:31:ARG:HD3	1:A:31:ARG:O	2.11	0.49
1:A:140:LYS:NZ	4:A:289:HOH:O	2.45	0.49
1:A:89:LEU:HD12	1:A:142:ILE:HD11	1.96	0.47
1:A:91:GLN:HA	1:A:149:LEU:HD13	1.96	0.47
1:A:81:HIS:HD2	4:A:290:HOH:O	1.98	0.46
1:A:0:MET:C	1:A:1:VAL:CG2	2.84	0.46
1:A:87:LYS:CB	1:A:88:PRO:CD	2.94	0.46
1:A:41:GLU:HA	1:A:47:LYS:CG	2.48	0.43
1:A:98:LYS:CE	4:A:302:HOH:O	2.54	0.43
1:A:45:ARG:CD	4:A:310:HOH:O	2.67	0.42
1:A:83:GLU:H	1:A:83:GLU:CD	2.24	0.41
1:A:45:ARG:NE	4:A:310:HOH:O	2.52	0.41

Continued from previous page..

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-1 Atom-2		Clash overlap (Å)
4:A:309:HOH:O	4:A:309:HOH:O[2_655]	2.02	0.18

### 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	А	152/154~(99%)	143~(94%)	9~(6%)	0	100 100

There are no Ramachandran outliers to report.

#### 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.

Ι	Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
	1	А	126/126~(100%)	115 (91%)	11 (9%)	10 6

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	45	ARG
1	А	47	LYS
1	А	59	GLU
1	А	62	LYS
1	А	102	LYS
1	А	115	LEU
1	А	117	SER
1	А	133	LYS
1	А	142	ILE
1	А	145	LYS
1	А	152	GLN

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such side chains are listed below:

Mol	Chain	Res	Type
1	А	8	GLN
1	А	116	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 monosaccharides 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).

Mal	Mol Type Chain Res Li		Link	В	Bond lengths			Bond angles		
INIOI	туре	Unam	nes		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	HEM	А	154	1	41,50,50	1.96	13 (31%)	45,82,82	1.72	8 (17%)
2	SO4	А	156	-	4,4,4	0.63	0	6,6,6	0.44	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	А	154	1	-	5/12/54/54	-

Mol	Chain	Res	Type	Atoms	Ζ	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
3	А	154	HEM	CHB-C1B	4.96	1.47	1.35
3	А	154	HEM	CAA-C2A	4.39	1.58	1.52
3	А	154	HEM	C3C-C2C	-4.30	1.34	1.40
3	А	154	HEM	C4D-C3D	3.71	1.51	1.45
3	А	154	HEM	CBB-CAB	2.82	1.44	1.30
3	А	154	HEM	C1D-ND	2.72	1.44	1.38
3	А	154	HEM	C4A-NA	2.44	1.41	1.36
3	А	154	HEM	C1A-NA	2.28	1.40	1.36
3	А	154	HEM	C1B-NB	-2.22	1.36	1.40
3	А	154	HEM	C4D-ND	-2.19	1.36	1.40
3	А	154	HEM	O2D-CGD	-2.12	1.23	1.30
3	А	154	HEM	CBC-CAC	2.09	1.43	1.29
3	А	154	HEM	O2A-CGA	-2.07	1.23	1.30

All (13) bond length outliers are listed below:

All (8) bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Ζ	$Observed(^{o})$	$Ideal(^{o})$
3	А	154	HEM	CBA-CAA-C2A	4.42	120.16	112.62
3	А	154	HEM	C3B-C2B-C1B	-4.36	103.25	106.49
3	А	154	HEM	CHD-C1D-ND	3.13	127.83	124.43
3	А	154	HEM	C1D-C2D-C3D	3.09	110.21	106.96
3	А	154	HEM	CHB-C1B-NB	-2.66	121.10	124.38
3	А	154	HEM	C4B-CHC-C1C	2.64	126.04	122.56
3	А	154	HEM	C2B-C1B-NB	2.61	112.93	109.84
3	А	154	HEM	CMA-C3A-C4A	-2.29	124.94	128.46

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	А	154	HEM	CAA-CBA-CGA-O2A
3	А	154	HEM	C2A-CAA-CBA-CGA
3	А	154	HEM	CAD-CBD-CGD-O1D
3	А	154	HEM	CAD-CBD-CGD-O2D
3	А	154	HEM	CAA-CBA-CGA-O1A

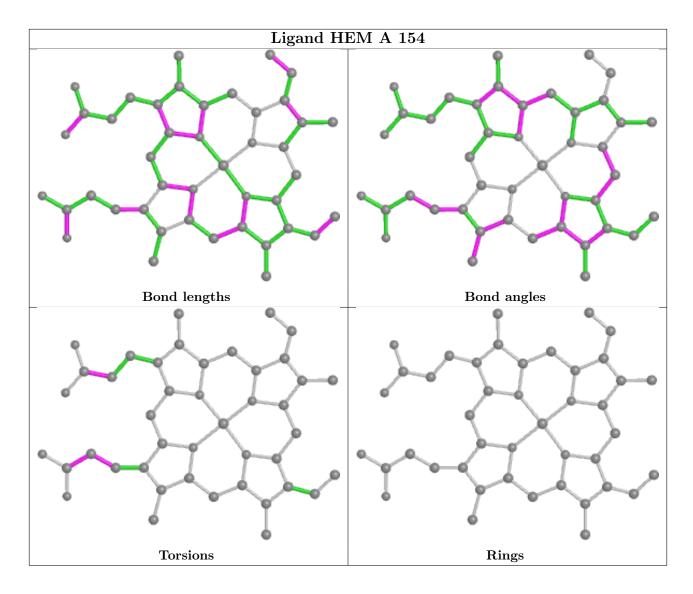
There are no ring outliers.

No monomer is involved in short contacts.

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.







## 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,  $95^{th}$  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	А	154/154~(100%)	-1.26	0 100 100	11, 18, 32, 50	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 monosaccharides in this entry.

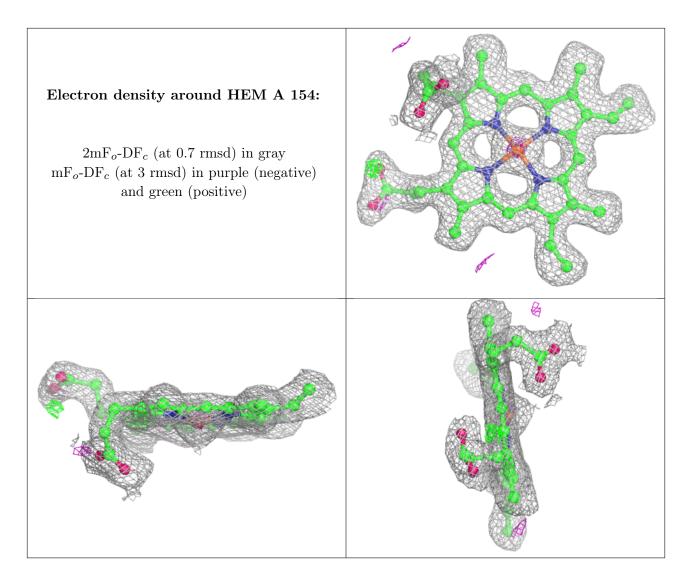
### 6.4 Ligands (i)

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,  $95^{th}$  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
2	SO4	А	156	5/5	0.98	0.08	48,48,49,49	0
3	HEM	А	154	43/43	0.99	0.05	12,14,24,30	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.





## 6.5 Other polymers (i)

There are no such residues in this entry.

