W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Feb 12, 2017 – 11:13 pm GMT

PDB ID : 1S5X
Title : The crystal structure of Trematomus bernacchii hemoglobin oxidized by air
Authors : Vitagliano, L.; Bonomi, G.; Riccio, A.; di Prisco, G.; Smulevich, G.; Mazzarella, L.
Deposited on : 2004-01-22
Resolution : 2.40 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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 2.40 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	Percentile Ranks	Value
Clashscore		7
Ramachandran outliers		0.4%
Sidechain outliers		2.2%

Worse .... *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
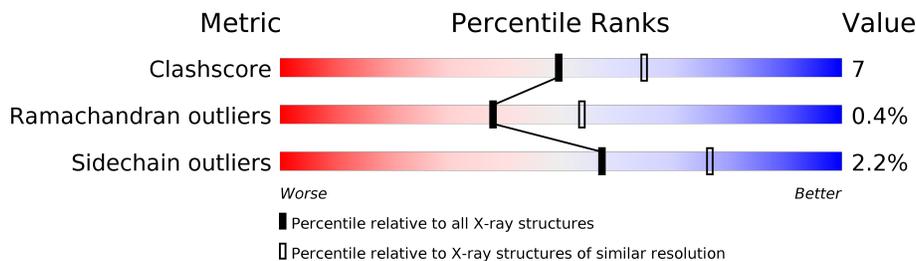
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	143	84% 15% •
2	B	146	77% 14% • 7%

Full wwPDB X-ray Structure Validation Report

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 2267 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 Hemoglobin alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	143	Total 1104	C 710	N 190	O 199	S 5	0	0	0

- Molecule 2 is a protein called Hemoglobin beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	136	Total 1050	C 669	N 182	O 194	S 5	0	0	0

- Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).
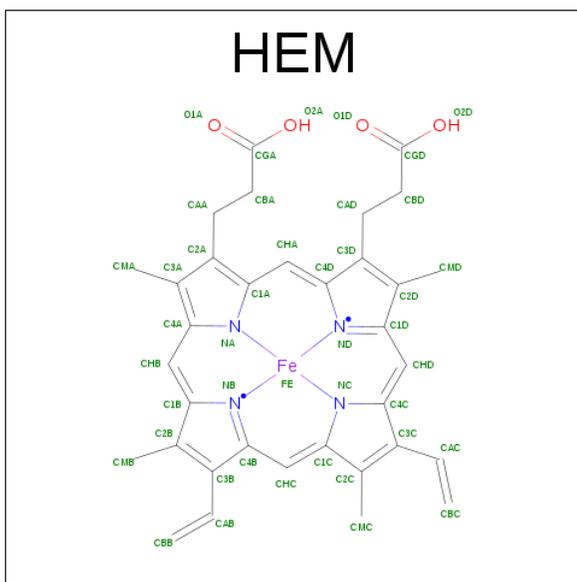
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
3	B	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

• Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	15	Total O 15 15		0	0
4	B	12	Total O 12 12		0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

Note EDS was not executed.

• Molecule 1: Hemoglobin alpha chain
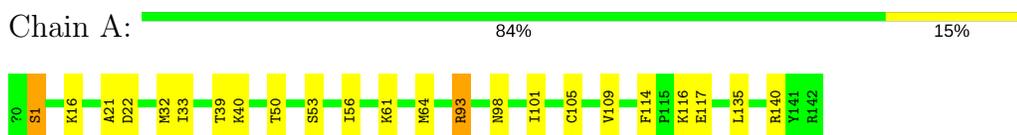
Chain A:

• Molecule 2: Hemoglobin beta chain

Chain B:
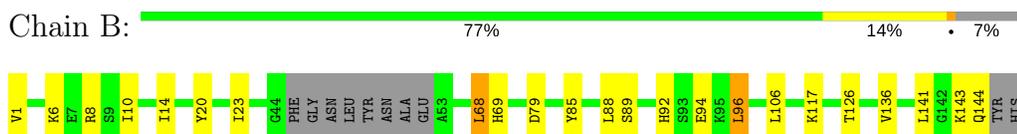
4 Data and refinement statistics ⓘ

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	108.52Å 65.09Å 55.75Å 90.00° 113.48° 90.00°	Depositor
Resolution (Å)	20.00 – 2.40	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-2.40)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.190 , 0.233	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2267	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	0.46	0/1127	0.68	1/1523 (0.1%)
2	B	0.42	0/1071	0.56	0/1449
All	All	0.44	0/2198	0.63	1/2972 (0.0%)

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	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1	SER	O-C-N	-11.73	103.93	122.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1	SER	Mainchain

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within

the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1104	0	1136	15	0
2	B	1050	0	1047	15	0
3	A	43	0	30	0	0
3	B	43	0	30	2	0
4	A	15	0	0	1	0
4	B	12	0	0	1	0
All	All	2267	0	2243	30	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 7.

All (30) 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:93:ARG:HH11	1:A:93:ARG:HB3	1.43	0.82
2:B:23:ILE:HD11	2:B:117:LYS:HD2	1.62	0.80
1:A:140:ARG:HB3	1:A:140:ARG:HH11	1.53	0.72
2:B:89:SER:OG	2:B:144:GLN:HB2	1.91	0.70
2:B:20:TYR:HA	2:B:68:LEU:HD23	1.73	0.70
1:A:140:ARG:HB3	1:A:140:ARG:NH1	2.12	0.65
1:A:116:LYS:HD2	1:A:117:GLU:OE2	1.97	0.65
1:A:22:ASP:OD1	1:A:61:LYS:HE3	2.05	0.56
2:B:1:VAL:HG11	2:B:136:VAL:HG22	1.88	0.56
2:B:8:ARG:HH12	2:B:79:ASP:CG	2.10	0.54
1:A:93:ARG:NH1	1:A:93:ARG:HB3	2.17	0.53
2:B:106:LEU:HD23	3:B:400:HEM:HBB2	1.90	0.52
1:A:105:CYS:O	1:A:109:VAL:HG23	2.12	0.49
2:B:69:HIS:HE1	4:B:516:HOH:O	1.93	0.49
2:B:85:TYR:HD2	2:B:88:LEU:HD12	1.79	0.47
1:A:116:LYS:HB3	4:A:502:HOH:O	2.15	0.47
2:B:141:LEU:CD1	3:B:400:HEM:HAB	2.45	0.46
2:B:89:SER:CB	2:B:144:GLN:HB2	2.46	0.46
1:A:33:ILE:HG23	1:A:40:LYS:HG2	1.97	0.45
2:B:6:LYS:O	2:B:10:ILE:HG13	2.16	0.44
1:A:114:PHE:HB3	1:A:117:GLU:OE1	2.17	0.44
1:A:50:THR:O	1:A:56:ILE:HD12	2.18	0.44
2:B:14:ILE:HD11	2:B:126:THR:HG23	2.00	0.43
2:B:92:HIS:HA	2:B:96:LEU:HB2	2.01	0.43
1:A:98:ASN:HD22	1:A:101:ILE:HD12	1.83	0.43
2:B:143:LYS:HE3	2:B:144:GLN:HE21	1.85	0.41

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:21:ALA:O	1:A:64:MET:HG3	2.20	0.41
1:A:32:MET:HE3	1:A:39:THR:HB	2.03	0.41
2:B:94:GLU:HA	2:B:94:GLU:OE2	2.20	0.40
1:A:140:ARG:HH11	1:A:140:ARG:CB	2.30	0.40

There are no symmetry-related clashes.

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.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	141/143 (99%)	137 (97%)	3 (2%)	1 (1%)	25	37
2	B	132/146 (90%)	127 (96%)	5 (4%)	0	100	100
All	All	273/289 (94%)	264 (97%)	8 (3%)	1 (0%)	38	54

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	53	SER

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	119/119 (100%)	116 (98%)	3 (2%)	53	73

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	112/120 (93%)	110 (98%)	2 (2%)	64	81
All	All	231/239 (97%)	226 (98%)	5 (2%)	57	76

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	16	LYS
1	A	93	ARG
1	A	135	LEU
2	B	68	LEU
2	B	96	LEU

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

Mol	Chain	Res	Type
1	A	98	ASN
1	A	103	ASN
2	B	41	HIS
2	B	69	HIS
2	B	144	GLN

5.3.3   RNA ⓘ

There are no RNA molecules 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 ⓘ

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	200	1,4	28,50,50	1.87	8 (28%)	17,82,82	2.14	6 (35%)
3	HEM	B	400	2	28,50,50	1.93	7 (25%)	17,82,82	2.27	7 (41%)

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	200	1,4	-	0/6/54/54	0/0/8/8
3	HEM	B	400	2	-	0/6/54/54	0/0/8/8

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	200	HEM	C3B-C2B	-3.35	1.35	1.40
3	A	200	HEM	C4B-NB	2.26	1.41	1.36
3	A	200	HEM	C1A-NA	2.33	1.41	1.36
3	B	400	HEM	C3B-CAB	2.61	1.53	1.47
3	A	200	HEM	C3B-CAB	2.63	1.53	1.47
3	A	200	HEM	C1B-NB	2.93	1.40	1.36
3	B	400	HEM	CBB-CAB	3.22	1.51	1.28
3	A	200	HEM	CBC-CAC	3.26	1.51	1.28
3	B	400	HEM	C1C-NC	3.38	1.40	1.36
3	B	400	HEM	C4C-NC	3.44	1.40	1.36
3	B	400	HEM	CMB-C2B	3.46	1.59	1.51
3	B	400	HEM	C3C-CAC	3.51	1.54	1.47
3	A	200	HEM	CBB-CAB	3.92	1.56	1.28
3	A	200	HEM	C3C-CAC	3.95	1.55	1.47
3	B	400	HEM	CBC-CAC	3.99	1.56	1.28

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	200	HEM	C1D-C2D-C3D	-5.18	103.39	107.00
3	A	200	HEM	CBD-CAD-C3D	-3.60	105.59	112.47
3	B	400	HEM	CMD-C2D-C1D	-2.82	124.13	128.46
3	B	400	HEM	C1D-C2D-C3D	-2.53	105.24	107.00
3	A	200	HEM	CMA-C3A-C4A	-2.22	125.06	128.46
3	B	400	HEM	C4A-C3A-C2A	-2.19	105.47	107.00
3	A	200	HEM	CMD-C2D-C3D	2.09	128.88	124.94
3	B	400	HEM	CMA-C3A-C2A	2.21	129.11	124.94
3	A	200	HEM	CMC-C2C-C3C	2.53	129.59	124.89
3	B	400	HEM	CMD-C2D-C3D	3.00	130.60	124.94
3	B	400	HEM	CAA-CBA-CGA	3.46	118.57	112.66
3	A	200	HEM	C3B-C4B-NB	3.52	113.76	109.21
3	B	400	HEM	CAD-CBD-CGD	4.87	120.97	112.66

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	400	HEM	2	0

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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.