



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 06:10 AM GMT

PDB ID : 2AA1  
Title : Crystal structure of the cathodic hemoglobin isolated from the Antarctic fish Trematomus Newnesi  
Authors : Mazzarella, L.; Bonomi, G.; Lubrano, M.C.; Merlino, A.; Riccio, A.; Vergara, A.; Vitagliano, L.; Verde, C.; Di Prisco, G.  
Deposited on : 2005-07-13  
Resolution : 1.80 Å(reported)

This is a full wwPDB validation 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>

---

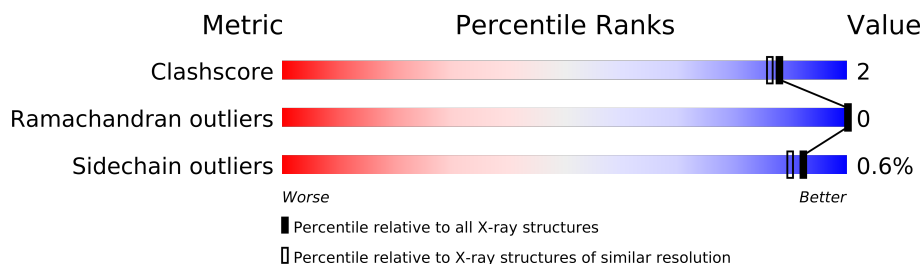
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 21963  
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 1.80 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	79885	4461 (1.80-1.80)
Ramachandran outliers	78287	4404 (1.80-1.80)
Sidechain outliers	78261	4403 (1.80-1.80)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	143	
1	C	143	
2	B	146	
2	D	146	

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 5081 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 Hemoglobin alpha-1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	143	Total	C	N	O	S	0	1	0
			1111	713	190	203	5			
1	C	143	Total	C	N	O	S	0	1	0
			1111	713	190	203	5			

- Molecule 2 is a protein called Hemoglobin beta-C chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	146	Total	C	N	O	S	0	2	0
			1159	743	195	216	5			
2	D	146	Total	C	N	O	S	0	2	0
			1159	743	195	216	5			

- 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
3	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
3	D	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	99	Total 99	O 99	0	0
4	B	80	Total 80	O 80	0	0
4	C	108	Total 108	O 108	0	0
4	D	82	Total 82	O 82	0	0

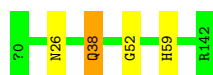
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

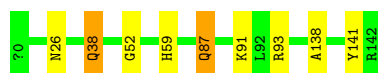
- Molecule 1: Hemoglobin alpha-1 chain

Chain A: 



- Molecule 1: Hemoglobin alpha-1 chain

Chain C: 



- Molecule 2: Hemoglobin beta-C chain

Chain B: 



- Molecule 2: Hemoglobin beta-C chain

Chain D: 



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.10Å 60.80Å 61.23Å 80.73° 81.36° 62.01°	Depositor
Resolution (Å)	20.00 – 1.80	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-1.80)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.186 , 0.220	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5081	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	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	0.31	0/1133	0.49	0/1530
1	C	0.31	0/1133	0.49	0/1530
2	B	0.33	0/1183	0.50	0/1603
2	D	0.33	0/1183	0.50	0/1603
All	All	0.32	0/4632	0.50	0/6266

There are no bond length outliers.

There are no bond angle outliers.

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	1111	0	1143	4	0
1	C	1111	0	1143	7	0
2	B	1159	0	1151	7	0
2	D	1159	0	1151	6	0
3	A	43	0	30	1	0
3	B	43	0	30	0	0
3	C	43	0	30	1	0
3	D	43	0	30	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	99	0	0	0	0
4	B	80	0	0	0	0
4	C	108	0	0	0	0
4	D	82	0	0	0	0
All	All	5081	0	4708	23	0

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 2.

All (23) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:38:GLN:HE21	1:A:38:GLN:H	1.30	0.80
1:C:38:GLN:H	1:C:38:GLN:HE21	1.31	0.75
2:B:20:TYR:HA	2:B:68:LEU:HD23	1.76	0.67
1:A:26:ASN:HD21	1:A:52:GLY:HA2	1.66	0.61
1:C:59:HIS:HE1	3:C:600:HEM:C4D	2.19	0.60
1:C:26:ASN:HD21	1:C:52:GLY:HA2	1.68	0.59
1:C:87:GLN:OE1	1:C:91:LYS:HD3	2.04	0.57
2:B:1:VAL:HG11	2:B:132:LYS:HE2	1.88	0.56
1:A:38:GLN:H	1:A:38:GLN:NE2	2.02	0.55
1:A:59:HIS:HE1	3:A:200:HEM:C4D	2.30	0.50
2:B:37:TRP:HA	1:C:93:ARG:HD2	1.94	0.49
2:B:82:THR:HA	2:B:140[B]:SER:OG	2.15	0.46
2:B:1:VAL:CG1	2:B:132:LYS:HE2	2.45	0.46
1:C:138:ALA:HA	1:C:141:TYR:CD1	2.51	0.45
2:D:82:THR:HG23	2:D:143:LYS:HD2	1.99	0.44
2:D:85:TYR:HB2	2:D:140[B]:SER:OG	2.18	0.43
2:D:41:TYR:CE2	2:D:98:VAL:HA	2.53	0.42
1:C:38:GLN:H	1:C:38:GLN:NE2	2.07	0.42
2:D:82:THR:HA	2:D:140[B]:SER:OG	2.20	0.41
2:D:18:LEU:HD22	2:D:23:VAL:HG21	2.02	0.41
2:B:74:ALA:HB2	2:B:85:TYR:HE2	1.85	0.41
2:B:41:TYR:CE2	2:B:98:VAL:HA	2.55	0.41
2:D:20:TYR:N	2:D:20:TYR:CD1	2.89	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	142/143 (99%)	140 (99%)	2 (1%)	0	100	100
1	C	142/143 (99%)	141 (99%)	1 (1%)	0	100	100
2	B	146/146 (100%)	145 (99%)	1 (1%)	0	100	100
2	D	146/146 (100%)	144 (99%)	2 (1%)	0	100	100
All	All	576/578 (100%)	570 (99%)	6 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	122/121 (101%)	121 (99%)	1 (1%)	89	85
1	C	122/121 (101%)	120 (98%)	2 (2%)	75	63
2	B	123/121 (102%)	123 (100%)	0	100	100
2	D	123/121 (102%)	123 (100%)	0	100	100
All	All	490/484 (101%)	487 (99%)	3 (1%)	92	90

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	GLN
1	C	38	GLN
1	C	87	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	26	ASN
1	A	38	GLN
1	A	87	GLN
2	B	63	HIS
1	C	26	ASN
1	C	38	GLN
1	C	59	HIS

### 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 ⓘ

4 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	49,50,50	2.10	18 (36%)	46,82,82	1.35	7 (15%)
3	HEM	B	400	2	49,50,50	2.21	15 (30%)	46,82,82	1.19	3 (6%)
3	HEM	C	600	1	49,50,50	2.63	14 (28%)	46,82,82	1.21	3 (6%)
3	HEM	D	800	2	49,50,50	2.44	15 (30%)	46,82,82	1.12	4 (8%)

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	-	0/14/114/114	0/0/8/8
3	HEM	B	400	2	-	0/14/114/114	0/0/8/8
3	HEM	C	600	1	-	0/14/114/114	0/0/8/8
3	HEM	D	800	2	-	0/14/114/114	0/0/8/8

All (62) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	600	HEM	C3D-C4D	11.45	1.47	1.44
3	D	800	HEM	C2B-C1B	8.36	1.46	1.44
3	B	400	HEM	C2D-C1D	8.00	1.46	1.44
3	D	800	HEM	C3D-C4D	7.05	1.46	1.44
3	C	600	HEM	C2D-C1D	6.88	1.46	1.44
3	A	200	HEM	C2B-C1B	5.65	1.46	1.44
3	B	400	HEM	C2B-C1B	5.28	1.45	1.44
3	A	200	HEM	C3D-C2D	-5.15	1.34	1.43
3	D	800	HEM	C2D-C1D	4.78	1.45	1.44
3	D	800	HEM	C3D-C2D	-4.58	1.35	1.43
3	C	600	HEM	C3B-C2B	-4.54	1.35	1.43
3	D	800	HEM	C4A-C3A	4.43	1.45	1.40
3	C	600	HEM	C4A-C3A	4.25	1.45	1.40
3	B	400	HEM	C3D-C2D	-4.23	1.36	1.43
3	A	200	HEM	C4A-C3A	4.19	1.45	1.40
3	C	600	HEM	C3D-C2D	-4.11	1.36	1.43
3	B	400	HEM	C4A-C3A	4.10	1.45	1.40
3	A	200	HEM	FE-NC	3.97	2.12	1.97
3	C	600	HEM	C3C-C2C	-3.92	1.36	1.43
3	B	400	HEM	C3B-C2B	-3.88	1.37	1.43
3	A	200	HEM	C3B-C2B	-3.75	1.37	1.43
3	A	200	HEM	C3C-C2C	-3.69	1.37	1.43
3	D	800	HEM	C3C-C2C	-3.65	1.37	1.43
3	D	800	HEM	CHA-C4D	3.61	1.41	1.35
3	A	200	HEM	C3D-C4D	3.51	1.45	1.44
3	D	800	HEM	C3B-C2B	-3.35	1.37	1.43
3	D	800	HEM	FE-NA	3.27	2.06	1.92
3	B	400	HEM	C3C-C2C	-3.20	1.38	1.43
3	C	600	HEM	FE-NB	3.14	2.09	1.97
3	B	400	HEM	CBC-CAC	3.08	1.46	1.28
3	D	800	HEM	CBC-CAC	2.98	1.46	1.28
3	D	800	HEM	CBB-CAB	2.96	1.46	1.28
3	C	600	HEM	FE-NA	2.92	2.04	1.92

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	400	HEM	FE-NC	2.92	2.08	1.97
3	B	400	HEM	CHA-C4D	2.90	1.40	1.35
3	B	400	HEM	CBB-CAB	2.89	1.45	1.28
3	A	200	HEM	CBC-CAC	2.85	1.45	1.28
3	A	200	HEM	FE-NA	2.85	2.04	1.92
3	A	200	HEM	CBB-CAB	2.85	1.45	1.28
3	A	200	HEM	CHA-C4D	2.80	1.39	1.35
3	C	600	HEM	CBC-CAC	2.76	1.44	1.28
3	C	600	HEM	CHA-C4D	2.61	1.39	1.35
3	C	600	HEM	CBB-CAB	2.56	1.43	1.28
3	A	200	HEM	CHC-C1C	2.47	1.40	1.36
3	B	400	HEM	FE-NA	2.41	2.02	1.92
3	C	600	HEM	C2B-C1B	2.41	1.45	1.44
3	B	400	HEM	C3D-C4D	2.40	1.45	1.44
3	A	200	HEM	C3B-C4B	2.34	1.47	1.44
3	A	200	HEM	CMB-C2B	2.33	1.54	1.47
3	B	400	HEM	CMB-C2B	2.33	1.54	1.47
3	C	600	HEM	CHC-C1C	2.32	1.40	1.36
3	D	800	HEM	CMB-C2B	2.30	1.54	1.47
3	A	200	HEM	C2D-C1D	2.29	1.45	1.44
3	D	800	HEM	CMD-C2D	2.15	1.54	1.47
3	C	600	HEM	CMB-C2B	2.12	1.54	1.47
3	D	800	HEM	FE-ND	2.09	2.05	1.97
3	D	800	HEM	CMC-C2C	2.04	1.53	1.47
3	B	400	HEM	CHD-C4C	2.03	1.39	1.36
3	A	200	HEM	CMD-C2D	2.03	1.53	1.47
3	A	200	HEM	CHB-C1B	2.01	1.38	1.35
3	A	200	HEM	CMC-C2C	2.01	1.53	1.47
3	B	400	HEM	CHB-C1B	2.00	1.38	1.35

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	200	HEM	C3B-C4B-NB	-4.53	110.76	114.00
3	C	600	HEM	C3B-C4B-NB	-4.31	110.92	114.00
3	B	400	HEM	C3B-C4B-NB	-3.95	111.17	114.00
3	D	800	HEM	C3B-C4B-NB	-3.18	111.72	114.00
3	A	200	HEM	CBD-CAD-C3D	-3.00	107.83	114.37
3	C	600	HEM	CAD-C3D-C4D	2.90	129.75	124.53
3	D	800	HEM	C3A-C4A-NA	2.86	111.57	109.41
3	A	200	HEM	CHD-C4C-NC	-2.79	122.31	124.73
3	A	200	HEM	CBA-CAA-C2A	2.72	117.49	112.69

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	400	HEM	C3A-C4A-NA	2.67	111.42	109.41
3	C	600	HEM	C3A-C4A-NA	2.59	111.36	109.41
3	A	200	HEM	CAD-C3D-C4D	2.30	128.67	124.53
3	D	800	HEM	CBA-CAA-C2A	-2.25	108.73	112.69
3	A	200	HEM	CAD-C3D-C2D	-2.11	122.55	127.25
3	A	200	HEM	C3A-C4A-NA	2.11	111.00	109.41
3	D	800	HEM	C4A-C3A-C2A	-2.10	105.53	107.00
3	B	400	HEM	CHD-C4C-NC	-2.07	122.93	124.73

There are no chirality outliers.

There are no torsion outliers.

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 ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.