



# wwPDB X-ray Structure Validation Summary Report

Oct 20, 2014 – 05:23 PM EDT

PDB ID : 2DHB  
Title : THREE DIMENSIONAL FOURIER SYNTHESIS OF HORSE DEOXY-HAEMOGLOBIN AT 2.8 ANGSTROMS RESOLUTION  
Authors : Perutz et al., M.F.  
Deposited on : 1973-11-01  
Resolution : 2.80 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>

---

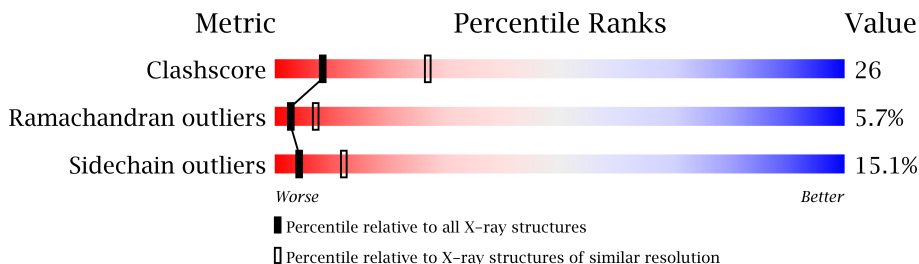
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 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) : stable24103

# 1 Overall quality at a glance

The reported resolution of this entry is 2.80 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	2295 (2.80-2.80)
Ramachandran outliers	78287	2252 (2.80-2.80)
Sidechain outliers	78261	2254 (2.80-2.80)

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	141	
2	B	146	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2289 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 (DEOXY) (ALPHA CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	141	Total	C	N	O	S	0	0	0
			1069	684	187	196	2			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	63	ALA	GLY	CONFLICT	UNP P01958
A	65	GLY	ALA	CONFLICT	UNP P01958
A	82	ASP	ASN	CONFLICT	UNP P01958
A	85	ASN	ASP	CONFLICT	UNP P01958

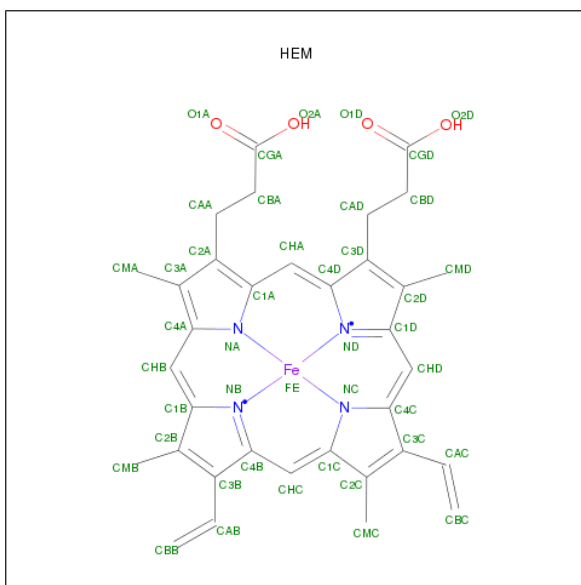
- Molecule 2 is a protein called HEMOGLOBIN (DEOXY) (BETA CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	146	Total	C	N	O	S	0	0	0
			1132	726	200	204	2			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	111	ALA	VAL	CONFLICT	UNP P02062
B	112	LEU	VAL	CONFLICT	UNP P02062
B	114	VAL	LEU	CONFLICT	UNP P02062

- 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.

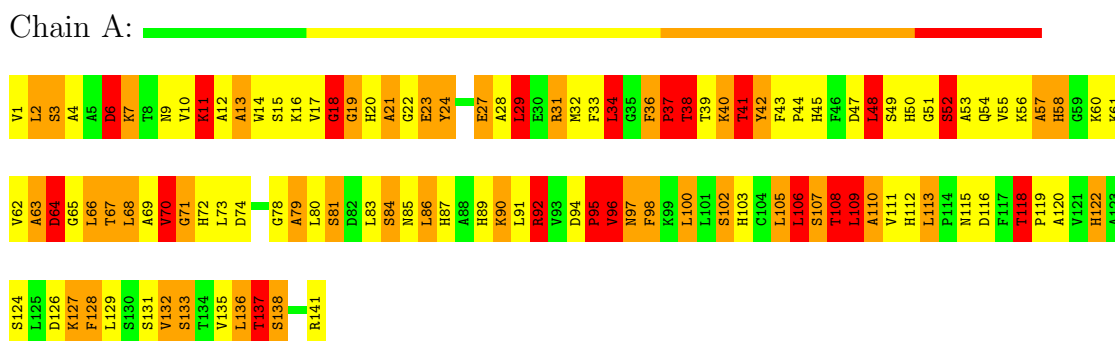
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total O 2 2	0	0

### 3 Residue-property plots

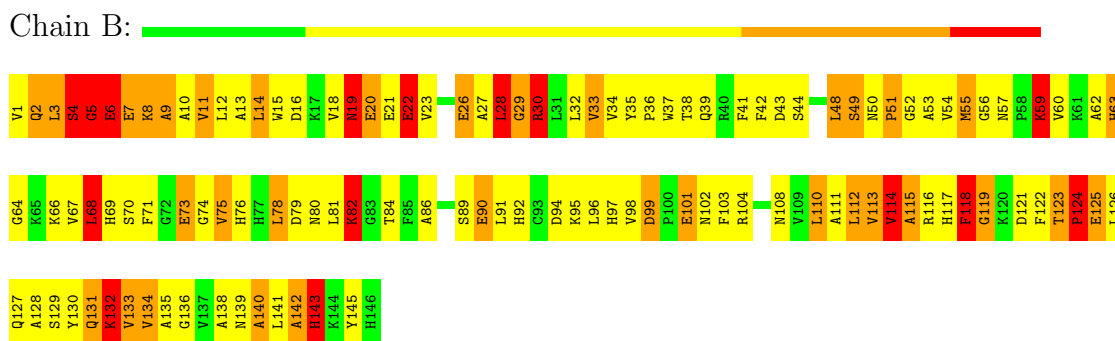
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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 (DEOXY) (ALPHA CHAIN)



#### • Molecule 2: HEMOGLOBIN (DEOXY) (BETA CHAIN)



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.96Å 81.70Å 92.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.80	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.80)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	unknown	Depositor
R, $R_{free}$	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2289	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	0.0	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

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	1.26	3/1096 (0.3%)	1.96	26/1487 (1.7%)
2	B	1.25	2/1160 (0.2%)	1.97	23/1569 (1.5%)
All	All	1.25	5/2256 (0.2%)	1.97	49/3056 (1.6%)

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	2	85
2	B	1	90
All	All	3	175

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	14	TRP	NE1-CE2	-6.85	1.28	1.37
2	B	15	TRP	NE1-CE2	-6.66	1.28	1.37
2	B	37	TRP	NE1-CE2	-6.63	1.28	1.37
1	A	3	SER	CB-OG	5.07	1.48	1.42
1	A	49	SER	CB-OG	5.07	1.48	1.42

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	109	LEU	N-CA-CB	9.99	130.38	110.40
2	B	5	GLY	C-N-CA	8.29	142.43	121.70
1	A	108	THR	N-CA-C	7.91	132.36	111.00
1	A	37	PRO	CA-N-CD	-7.32	101.25	111.50

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	108	THR	C-N-CA	7.06	139.35	121.70

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	108	THR	CA
1	A	109	LEU	CA
2	B	28	LEU	CA

5 of 175 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1	VAL	Mainchain
1	A	2	LEU	Mainchain
1	A	3	SER	Mainchain
1	A	4	ALA	Mainchain
1	A	6	ASP	Mainchain

## 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	1069	0	1073	69	52
2	B	1132	0	1120	64	0
3	A	43	0	30	6	0
3	B	43	0	30	2	0
4	A	2	0	0	0	0
All	All	2289	0	2253	119	52

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 26.

The worst 5 of 119 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:83:LEU:HD11	3:A:142:HEM:HMA3	1.50	0.92

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:78:LEU:HD21	2:B:133:VAL:HG22	1.49	0.91
1:A:29:LEU:HD11	1:A:58:HIS:HD2	1.39	0.86
2:B:11:VAL:HG11	2:B:133:VAL:HG21	1.61	0.82
1:A:95:PRO:HB3	1:A:137:THR:CG2	2.12	0.80

The worst 5 of 52 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:15:SER:OG	1:A:15:SER:OG[3_554]	0.49	1.71
1:A:19:GLY:CA	1:A:70:VAL:O[3_554]	0.55	1.65
1:A:67:THR:CB	1:A:67:THR:CG2[3_554]	1.06	1.14
1:A:19:GLY:CA	1:A:70:VAL:C[3_554]	1.14	1.06
1:A:67:THR:CG2	1:A:67:THR:CG2[3_554]	1.17	1.03

## 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	139/141 (99%)	113 (81%)	17 (12%)	9 (6%)	2	5
2	B	144/146 (99%)	108 (75%)	29 (20%)	7 (5%)	3	10
All	All	283/287 (99%)	221 (78%)	46 (16%)	16 (6%)	3	7

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	37	PRO
1	A	52	SER
1	A	109	LEU
2	B	5	GLY
2	B	19	ASN

### 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	115/115 (100%)	98 (85%)	17 (15%)	4	13
2	B	117/117 (100%)	99 (85%)	18 (15%)	4	12
All	All	232/232 (100%)	197 (85%)	35 (15%)	4	12

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	137	THR
2	B	4	SER
2	B	132	LYS
1	A	138	SER
2	B	2	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	103	HIS
2	B	76	HIS
1	A	122	HIS
1	A	97	ASN
2	B	63	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

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	142	1	42,50,50	3.54	16 (38%)	27,82,82	1.55	5 (18%)
3	HEM	B	147	2	42,50,50	3.42	16 (38%)	27,82,82	1.29	1 (3%)

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	142	1	-	0/14/114/114	0/0/8/8
3	HEM	B	147	2	-	0/14/114/114	0/0/8/8

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	142	HEM	C3C-C4C	-11.28	1.36	1.45
3	B	147	HEM	C1B-C2B	-9.56	1.37	1.45
3	A	142	HEM	C1B-C2B	-9.00	1.38	1.45
3	B	147	HEM	C3C-C2C	-7.26	1.39	1.45
3	B	147	HEM	C3C-C4C	-6.58	1.40	1.45

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	142	HEM	C3A-C4A-NA	-4.12	106.74	109.50
3	B	147	HEM	C3A-C4A-NA	-3.79	106.97	109.50
3	A	142	HEM	O2D-CGD-CBD	2.83	123.96	114.19
3	A	142	HEM	C4A-NA-C1A	2.74	111.35	107.93
3	A	142	HEM	CAD-CBD-CGD	-2.45	108.91	113.53

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.