



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 07:50 PM GMT

PDB ID : 1HGB
Title : HIGH RESOLUTION CRYSTAL STRUCTURES AND COMPARISONS OF T STATE DEOXYHAEMOGLOBIN AND TWO LIGANDED T-STATE HAEMOGLOBINS: T(ALPHA-OXY)HAEMOGLOBIN AND T(MET)HAEMOGLOBIN
Authors : Liddington, R.; Derewenda, Z.; Dodson, E.; Hubbard, R.; Dodson, G.
Deposited on : 1991-10-31
Resolution : 2.10 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

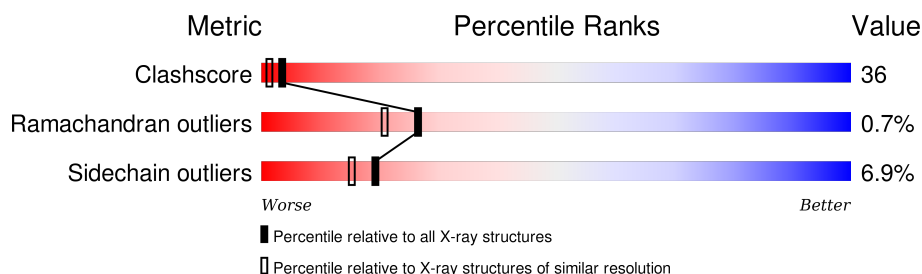
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)

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 $\leq 5\%$. 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	141	
1	C	141	
2	B	146	
2	D	146	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4762 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 (AQUO MET) (ALPHA CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	141	Total	C	N	O	S	0	0	0
			1069	685	187	194	3			
1	C	141	Total	C	N	O	S	0	0	0
			1069	685	187	194	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	146	Total	C	N	O	S	0	0	0
			1123	724	195	201	3			
2	D	146	Total	C	N	O	S	0	0	0
			1123	724	195	201	3			

- 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.

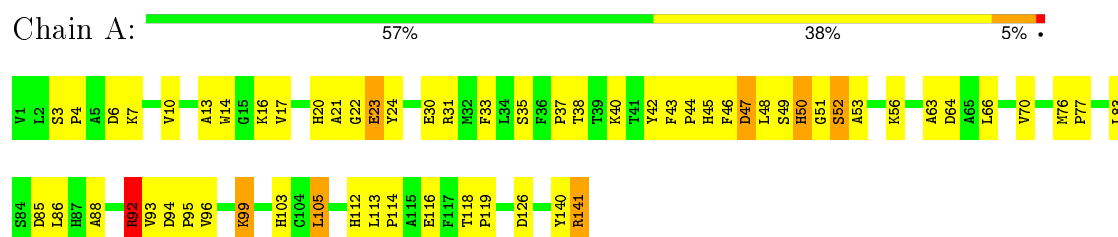
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	62	Total 62	O 62	0	0
4	B	50	Total 50	O 50	0	0
4	C	41	Total 41	O 41	0	0
4	D	53	Total 53	O 53	0	0

3 Residue-property plots

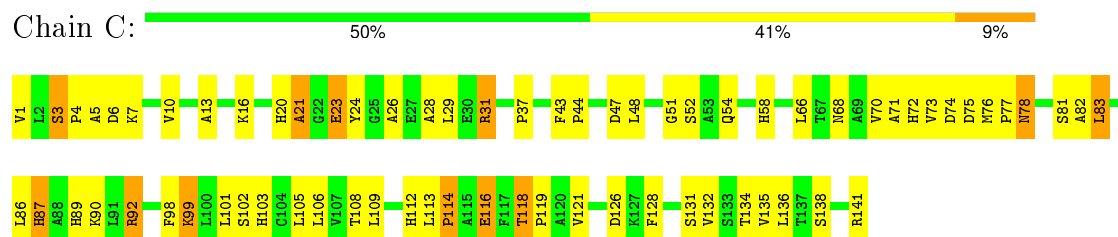
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

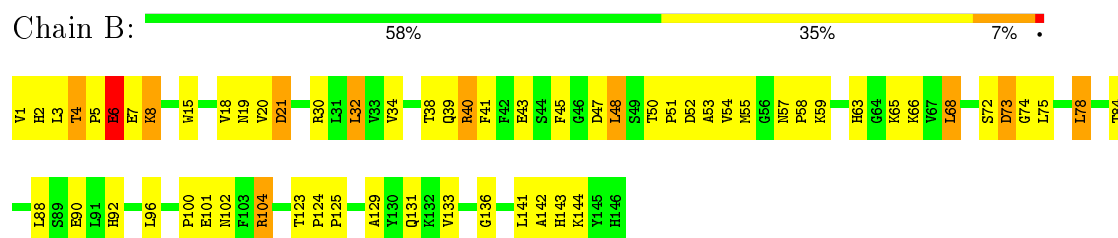
• Molecule 1: HEMOGLOBIN (AQUO MET) (ALPHA CHAIN)



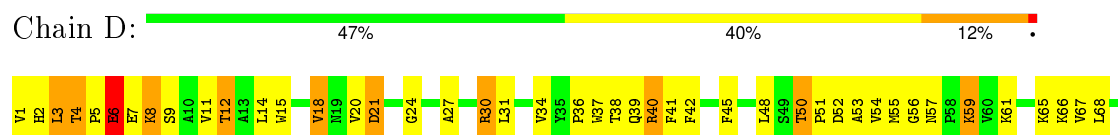
• Molecule 1: HEMOGLOBIN (AQUO MET) (ALPHA CHAIN)



• Molecule 2: HEMOGLOBIN (AQUO MET) (BETA CHAIN)



• Molecule 2: HEMOGLOBIN (AQUO MET) (BETA CHAIN)



D73	G74	H77	L78	D79	I80	L81	K82	G83	T84	F85	L86	S89	E90	D94	K95	L96	H97	Y98	D99	P100	F101	H102	F103	R104	L105	V109	H116	K120	T123	P124	P125	V126	Y130	Q131	V134	H143	K144	Y145	H146
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	95.78Å 97.78Å 65.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.10	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.10)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	0.210 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4762	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

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: 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	0.74	0/1097	1.54	12/1491 (0.8%)
1	C	0.76	0/1097	1.57	14/1491 (0.9%)
2	B	0.79	0/1153	1.52	9/1566 (0.6%)
2	D	0.76	0/1153	1.62	17/1566 (1.1%)
All	All	0.76	0/4500	1.57	52/6114 (0.9%)

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
1	C	0	2
2	B	0	1
All	All	0	4

There are no bond length outliers.

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	64	ASP	CB-CG-OD2	10.65	127.89	118.30
2	D	104	ARG	NE-CZ-NH1	10.31	125.45	120.30
1	C	141	ARG	NE-CZ-NH1	9.68	125.14	120.30
2	B	40	ARG	CD-NE-CZ	9.55	136.97	123.60
2	D	40	ARG	NE-CZ-NH1	9.53	125.06	120.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	92	ARG	Sidechain
2	B	104	ARG	Sidechain
1	C	31	ARG	Sidechain
1	C	92	ARG	Sidechain

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	1069	0	1073	57	2
1	C	1069	0	1073	85	2
2	B	1123	0	1118	78	1
2	D	1123	0	1118	113	0
3	A	43	0	30	1	0
3	B	43	0	30	2	0
3	C	43	0	30	0	0
3	D	43	0	30	4	0
4	A	62	0	0	6	3
4	B	50	0	0	24	5
4	C	41	0	0	38	1
4	D	53	0	0	20	2
All	All	4762	0	4502	328	8

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 36.

The worst 5 of 328 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)
2:D:15:TRP:CD1	4:D:198:HOH:O	1.71	1.35
2:B:6:GLU:CD	4:B:176:HOH:O	1.66	1.33
1:C:101:LEU:HD12	4:C:201:HOH:O	1.29	1.29
2:D:18:VAL:HG13	4:D:225:HOH:O	1.16	1.29
2:B:141:LEU:CA	4:B:170:HOH:O	1.75	1.27

The worst 5 of 8 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)
4:B:173:HOH:O	4:C:165:HOH:O[3_555]	0.60	1.60
1:C:99:LYS:NZ	4:A:200:HOH:O[4_456]	1.27	0.93
1:C:99:LYS:CE	4:A:200:HOH:O[4_456]	1.74	0.46
1:A:23:GLU:CB	4:B:186:HOH:O[3_555]	1.95	0.25
1:A:23:GLU:CG	4:B:186:HOH:O[3_555]	2.00	0.20

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	139/141 (99%)	131 (94%)	7 (5%)	1 (1%)	26	21
1	C	139/141 (99%)	131 (94%)	5 (4%)	3 (2%)	8	3
2	B	144/146 (99%)	137 (95%)	7 (5%)	0	100	100
2	D	144/146 (99%)	138 (96%)	6 (4%)	0	100	100
All	All	566/574 (99%)	537 (95%)	25 (4%)	4 (1%)	26	21

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	21	ALA
1	C	73	VAL
1	A	52	SER
1	C	114	PRO

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	113/113 (100%)	109 (96%)	4 (4%)	43	44
1	C	113/113 (100%)	109 (96%)	4 (4%)	43	44
2	B	118/118 (100%)	110 (93%)	8 (7%)	20	16
2	D	118/118 (100%)	102 (86%)	16 (14%)	5	2
All	All	462/462 (100%)	430 (93%)	32 (7%)	19	15

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	99	LYS
2	D	4	THR
2	D	95	LYS
2	D	3	LEU
2	D	6	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	80	ASN
2	B	102	ASN
1	C	103	HIS
2	B	77	HIS
1	C	97	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 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	142	1,4	30,50,50	2.37	7 (23%)	24,82,82	3.07	12 (50%)
3	HEM	B	147	2,4	30,50,50	2.53	10 (33%)	24,82,82	2.94	13 (54%)
3	HEM	C	142	1,4	30,50,50	2.46	9 (30%)	24,82,82	3.02	11 (45%)
3	HEM	D	147	2,4	30,50,50	2.64	10 (33%)	24,82,82	2.81	11 (45%)

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	1,4	-	0/10/54/54	0/0/8/8
3	HEM	B	147	2,4	-	0/10/54/54	0/0/8/8
3	HEM	C	142	1,4	-	0/10/54/54	0/0/8/8
3	HEM	D	147	2,4	-	0/10/54/54	0/0/8/8

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	147	HEM	C3B-C4B	-8.40	1.44	1.51
3	C	142	HEM	C3B-C4B	-7.05	1.45	1.51
3	D	147	HEM	C2D-C3D	-6.99	1.33	1.54
3	A	142	HEM	C2D-C3D	-6.84	1.34	1.54
3	B	147	HEM	C2D-C3D	-6.82	1.34	1.54

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	142	HEM	CAA-C2A-C1A	-5.59	120.93	127.01
3	D	147	HEM	CAA-C2A-C1A	-4.97	121.62	127.01

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	147	HEM	CMA-C3A-C4A	-4.93	120.21	128.36
3	C	142	HEM	CMA-C3A-C4A	-4.78	120.46	128.36
3	A	142	HEM	C3C-CAC-CBC	-4.61	117.39	124.46

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	142	HEM	1	0
3	B	147	HEM	2	0
3	D	147	HEM	4	0

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](#)

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

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

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

6.4 Ligands [i](#)

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

6.5 Other polymers [i](#)

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