



# Full wwPDB/EMDatabank EM Map/Model Validation Report ⓘ

Oct 10, 2018 – 03:11 PM EDT

PDB ID : 5NI1  
EMDB ID: : EMD-3488  
Title : CryoEM structure of haemoglobin at 3.2 Å determined with the Volta phase plate  
Authors : Khoshouei, M.; Radjainia, M.; Bunker, R.; Baumeister, W.; Danev, R.  
Deposited on : 2017-03-22  
Resolution : 3.20 Å(reported)

This is a Full wwPDB/EMDatabank EM Map/Model Validation Report  
for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20031633

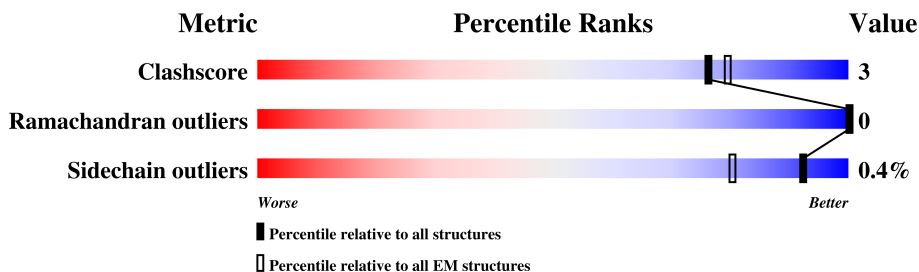
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.20 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)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	141	99% .
1	C	141	99% .
2	B	146	90% 10%
2	D	146	90% 10%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4579 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

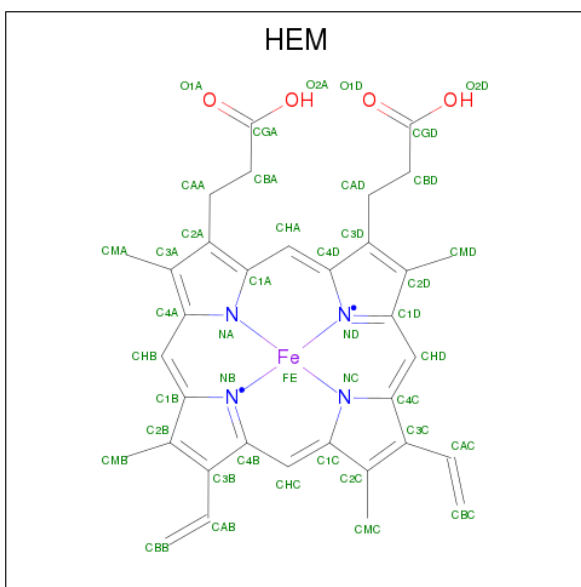
- Molecule 1 is a protein called Hemoglobin subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	141	Total	C	N	O	S	0	0
			1069	685	187	194	3		
1	C	141	Total	C	N	O	S	0	0
			1068	685	187	193	3		

- Molecule 2 is a protein called Hemoglobin subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	146	Total	C	N	O	S	0	0
			1123	724	195	201	3		
2	D	146	Total	C	N	O	S	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					AltConf
3	A	1	Total	C	Fe	N	O	0
			43	34	1	4	4	
3	B	1	Total	C	Fe	N	O	0
			43	34	1	4	4	
3	C	1	Total	C	Fe	N	O	0
			43	34	1	4	4	
3	D	1	Total	C	Fe	N	O	0
			43	34	1	4	4	

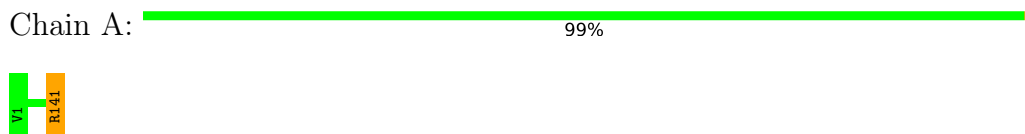
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		AltConf
4	A	7	Total	O	0
			7	7	
4	B	5	Total	O	0
			5	5	
4	C	7	Total	O	0
			7	7	
4	D	5	Total	O	0
			5	5	

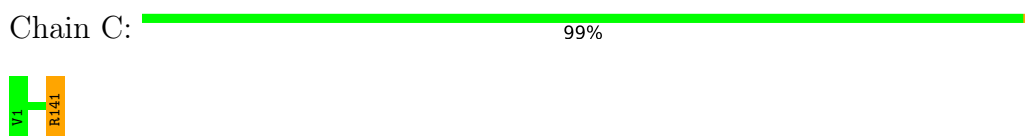
### 3 Residue-property plots [i](#)

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. 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. 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.

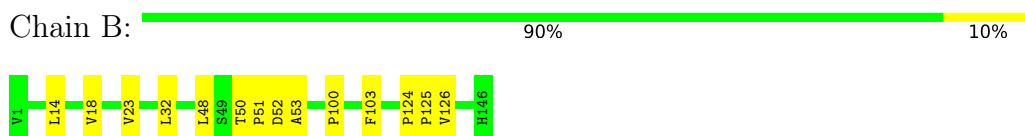
- Molecule 1: Hemoglobin subunit alpha



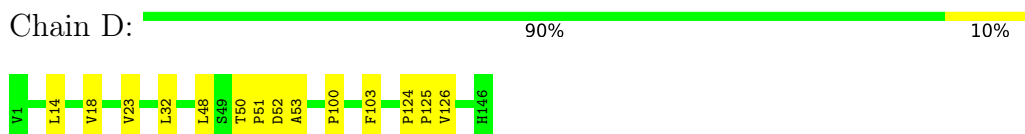
- Molecule 1: Hemoglobin subunit alpha



- Molecule 2: Hemoglobin subunit beta



- Molecule 2: Hemoglobin subunit beta



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	175300	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	40	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

## 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  >2	RMSZ	# Z  >2
1	A	0.48	0/1097	0.64	0/1491
1	C	0.48	0/1096	0.64	0/1490
2	B	0.46	0/1153	0.63	0/1566
2	D	0.46	0/1153	0.63	0/1566
All	All	0.47	0/4499	0.63	0/6113

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	A	1069	0	1073	4	0
1	C	1068	0	1070	4	0
2	B	1123	0	1118	8	0
2	D	1123	0	1118	8	0
3	A	43	0	30	1	0
3	B	43	0	30	2	0
3	C	43	0	30	1	0
3	D	43	0	30	2	0
4	A	7	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	5	0	0	0	0
4	C	7	0	0	0	0
4	D	5	0	0	0	0
All	All	4579	0	4499	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 3.

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:C:141:ARG:HG2	1:C:141:ARG:HH11	1.37	0.89
1:A:141:ARG:HH11	1:A:141:ARG:HG2	1.38	0.88
3:D:201:HEM:HBC2	3:D:201:HEM:HMC2	1.77	0.66
1:C:141:ARG:HG2	1:C:141:ARG:NH1	2.13	0.63
1:A:141:ARG:NH1	1:A:141:ARG:HG2	2.13	0.61
3:B:201:HEM:HMC1	3:B:201:HEM:HBC2	1.82	0.61
2:B:50:THR:O	2:B:51:PRO:C	2.43	0.56
1:A:141:ARG:HH11	1:A:141:ARG:CG	2.13	0.56
1:C:141:ARG:CG	1:C:141:ARG:HH11	2.13	0.53
1:A:141:ARG:NH1	1:A:141:ARG:CG	2.73	0.52
2:D:50:THR:O	2:D:51:PRO:C	2.44	0.51
1:C:141:ARG:CG	1:C:141:ARG:NH1	2.73	0.48
2:D:18:VAL:HG13	2:D:23:VAL:HG21	1.95	0.48
2:B:18:VAL:HG13	2:B:23:VAL:HG21	1.95	0.47
2:D:14:LEU:HD22	2:D:126:VAL:HG11	1.99	0.45
2:B:14:LEU:HD22	2:B:126:VAL:HG11	1.99	0.45
2:B:32:LEU:HD23	2:B:48:LEU:HD22	1.98	0.45
2:D:32:LEU:HD23	2:D:48:LEU:HD22	1.98	0.44
2:B:50:THR:O	2:B:52:ASP:N	2.52	0.43
2:B:124:PRO:HB2	2:B:125:PRO:HD3	2.01	0.43
2:B:100:PRO:HA	2:B:103:PHE:CD2	2.55	0.42
3:C:201:HEM:HBB2	3:C:201:HEM:HHC	2.01	0.42
2:D:124:PRO:HB2	2:D:125:PRO:HD3	2.01	0.42
2:D:100:PRO:HA	2:D:103:PHE:CD2	2.54	0.42
2:D:50:THR:O	2:D:52:ASP:N	2.53	0.42
3:D:201:HEM:HHC	3:D:201:HEM:HBB2	2.02	0.42
2:B:50:THR:O	2:B:53:ALA:N	2.50	0.42
2:D:50:THR:O	2:D:53:ALA:N	2.50	0.41
3:A:201:HEM:HHC	3:A:201:HEM:HBB2	2.02	0.41
3:B:201:HEM:HBB2	3:B:201:HEM:HHC	2.02	0.41



There are no symmetry-related clashes.

## 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 PDB entries followed by that with respect to all EM entries.

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%)	8 (6%)	0	100	100
1	C	139/141 (99%)	131 (94%)	8 (6%)	0	100	100
2	B	144/146 (99%)	138 (96%)	6 (4%)	0	100	100
2	D	144/146 (99%)	139 (96%)	5 (4%)	0	100	100
All	All	566/574 (99%)	539 (95%)	27 (5%)	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 PDB entries followed by that with respect to all EM entries.

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%)	112 (99%)	1 (1%)	81	93
1	C	112/113 (99%)	111 (99%)	1 (1%)	81	93
2	B	118/118 (100%)	118 (100%)	0	100	100
2	D	118/118 (100%)	118 (100%)	0	100	100
All	All	461/462 (100%)	459 (100%)	2 (0%)	92	96

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	141	ARG
1	C	141	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	9	ASN
2	B	63	HIS
1	C	9	ASN
2	D	63	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 carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

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	201	1	27,50,50	1.34	1 (3%)	17,82,82	1.45	3 (17%)
3	HEM	B	201	2	27,50,50	0.97	2 (7%)	17,82,82	1.24	1 (5%)
3	HEM	C	201	1	27,50,50	1.34	1 (3%)	17,82,82	1.46	3 (17%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	HEM	D	201	2	27,50,50	0.98	2 (7%)	17,82,82	1.22	1 (5%)

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	201	1	-	0/6/54/54	0/0/8/8
3	HEM	B	201	2	-	0/6/54/54	0/0/8/8
3	HEM	C	201	1	-	0/6/54/54	0/0/8/8
3	HEM	D	201	2	-	0/6/54/54	0/0/8/8

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	201	HEM	C3B-C2B	-5.14	1.33	1.40
3	A	201	HEM	C3B-C2B	-5.08	1.33	1.40
3	D	201	HEM	C3B-C2B	-3.31	1.35	1.40
3	B	201	HEM	C3B-C2B	-3.29	1.35	1.40
3	B	201	HEM	C4D-C3D	2.14	1.47	1.42
3	D	201	HEM	C4D-C3D	2.15	1.47	1.42

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	201	HEM	CAA-CBA-CGA	-2.70	108.05	112.66
3	D	201	HEM	CAA-CBA-CGA	-2.64	108.15	112.66
3	C	201	HEM	CAA-CBA-CGA	-2.61	108.20	112.66
3	A	201	HEM	CAA-CBA-CGA	-2.57	108.27	112.66
3	C	201	HEM	CMA-C3A-C4A	-2.24	125.02	128.46
3	A	201	HEM	CMA-C3A-C4A	-2.23	125.03	128.46
3	C	201	HEM	C3B-C4B-NB	-2.17	106.41	109.21
3	A	201	HEM	C3B-C4B-NB	-2.17	106.41	109.21

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	201	HEM	1	0
3	B	201	HEM	2	0
3	C	201	HEM	1	0
3	D	201	HEM	2	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.