



# wwPDB X-ray Structure Validation Summary Report

Jul 28, 2014 – 10:42 AM EDT

PDB ID : 3WCU  
Title : The structure of a deoxygenated 400 kda hemoglobin provides a more accurate description of the cooperative mechanism of giant hemoglobins: Deoxygenated form  
Authors : Numoto, N.; Nakagawa, T.; Ohara, R.; Hasegawa, T.; Kita, A.; Yoshida, T.; Maruyama, T.; Imai, K.; Fukumori, Y.; Miki, K.  
Deposited on : 2013-06-01  
Resolution : 2.90 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>

---

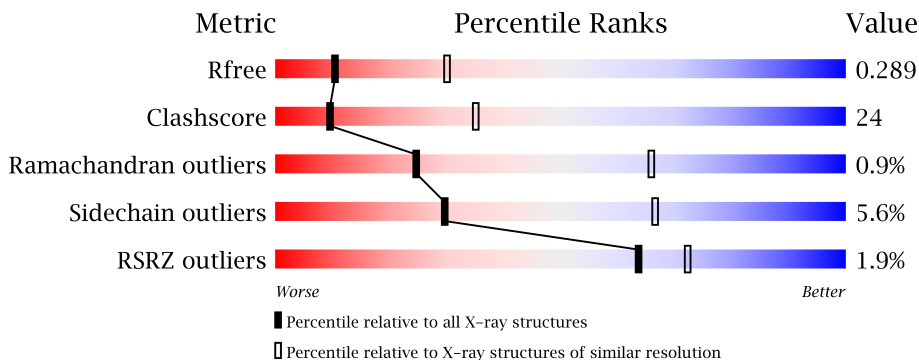
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	<b>FAILED</b>
Xtriage (Phenix)	:	dev-1439
EDS	:	stable23489
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable23489

# 1 Overall quality at a glance

The reported resolution of this entry is 2.90 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))
$R_{free}$	66092	1053 (2.90-2.90)
Clashscore	79885	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RSRZ outliers	66119	1054 (2.90-2.90)

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.

Mol	Chain	Length	Quality of chain
1	A	146	
1	E	146	
2	B	144	
2	F	144	
3	C	150	
3	G	150	
4	D	149	
4	H	149	

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 9626 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 A1 globin chain of giant V2 hemoglobin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	146	Total	C	N	O	S	0	0	0
			1151	736	198	209	8			
1	E	146	Total	C	N	O	S	0	0	0
			1151	736	198	209	8			

- Molecule 2 is a protein called A2 globin chain of giant V2 hemoglobin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	144	Total	C	N	O	S	0	0	0
			1120	697	206	210	7			
2	F	144	Total	C	N	O	S	0	0	0
			1120	697	206	210	7			

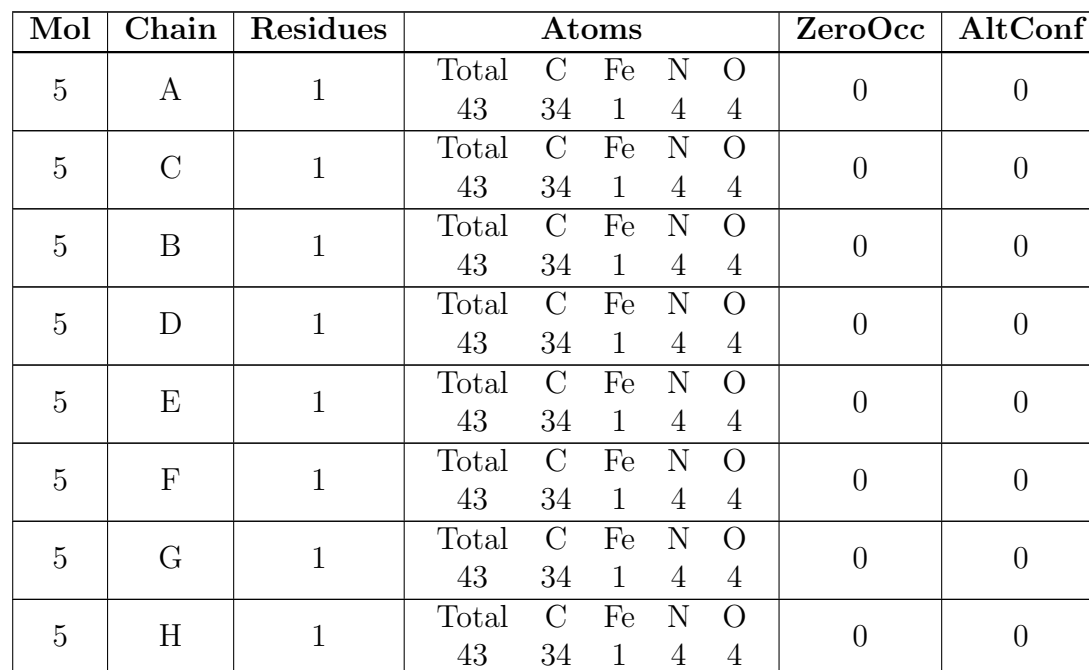
- Molecule 3 is a protein called B2 globin chain of giant V2 hemoglobin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	150	Total	C	N	O	S	0	0	0
			1158	718	211	221	8			
3	G	150	Total	C	N	O	S	0	0	0
			1158	718	211	221	8			

- Molecule 4 is a protein called B1 globin chain of giant V2 hemoglobin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	149	Total	C	N	O	S	0	0	0
			1140	727	196	210	7			
4	H	149	Total	C	N	O	S	0	0	0
			1140	727	196	210	7			

- Molecule 5 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: C<sub>34</sub>H<sub>32</sub>FeN<sub>4</sub>O<sub>4</sub>).



- | Mol | Chain | Residues | Atoms           | ZeroOcc | AltConf |
|-----|-------|----------|-----------------|---------|---------|
| 6   | B     | 1        | Total Ca<br>1 1 | 0       | 0       |
| 6   | F     | 1        | Total Ca<br>1 1 | 0       | 0       |

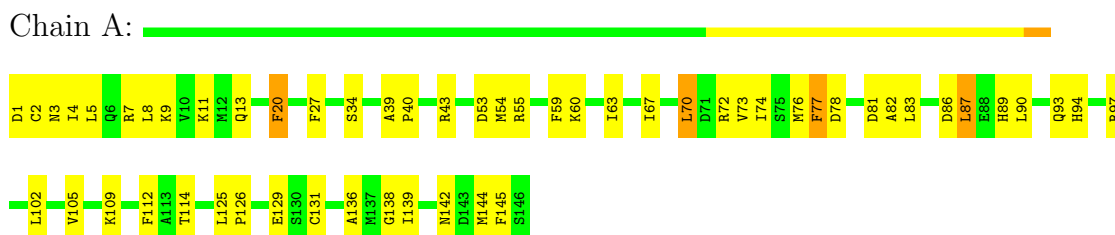
- 

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	D	6	Total 71	C 40	N 2	O 29	0	0
7	H	6	Total 71	C 40	N 2	O 29	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.

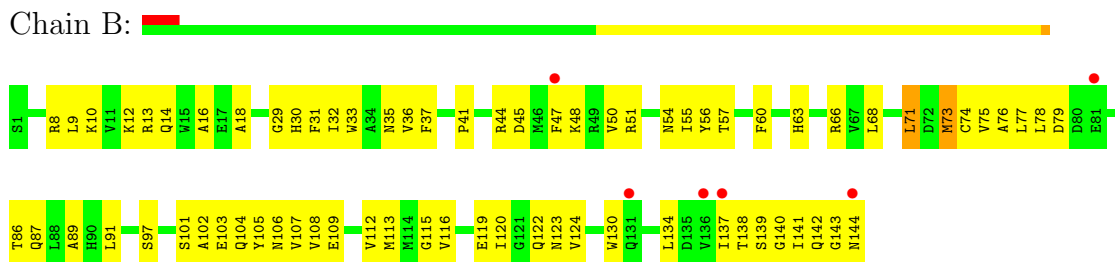
- Molecule 1: A1 globin chain of giant V2 hemoglobin



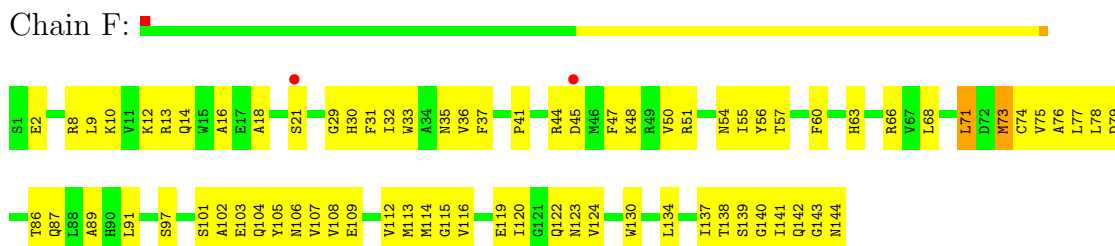
- Molecule 1: A1 globin chain of giant V2 hemoglobin



- Molecule 2: A2 globin chain of giant V2 hemoglobin

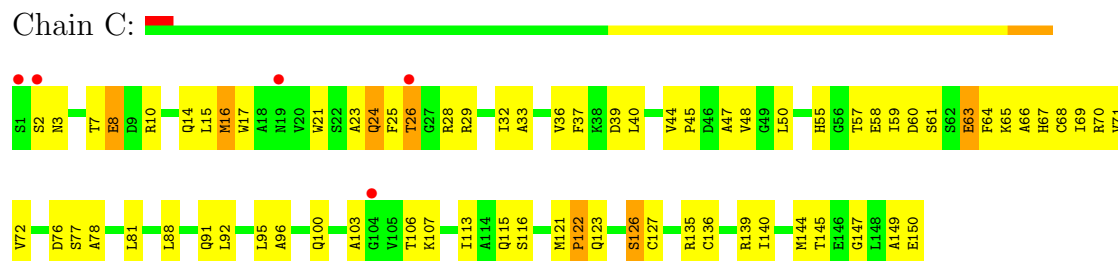


- Molecule 2: A2 globin chain of giant V2 hemoglobin



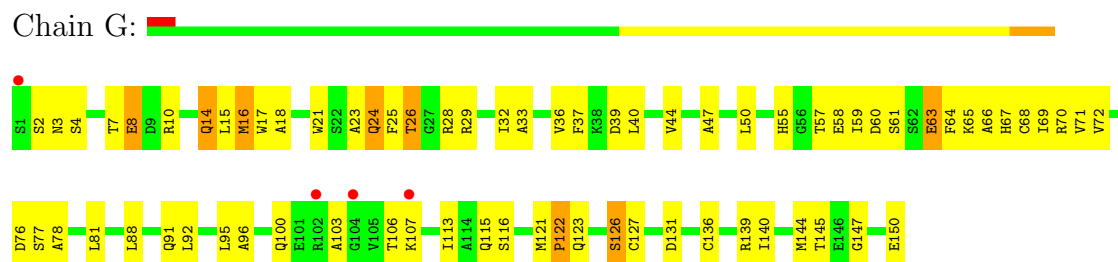
- Molecule 3: B2 globin chain of giant V2 hemoglobin

Chain C:



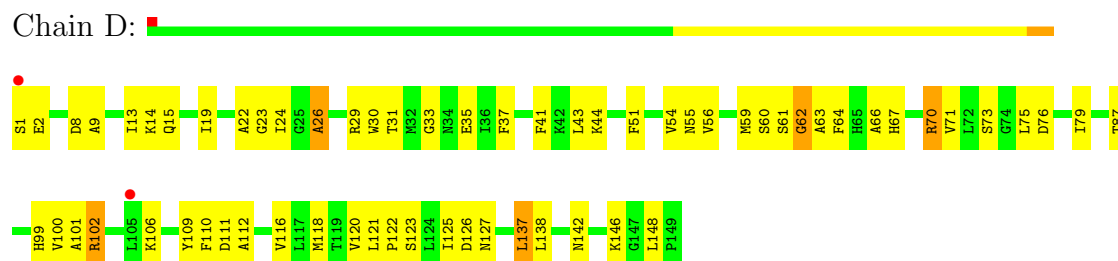
- Molecule 3: B2 globin chain of giant V2 hemoglobin

Chain G:



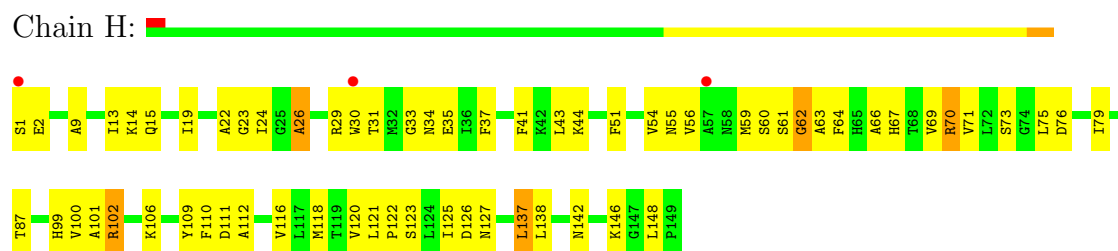
- Molecule 4: B1 globin chain of giant V2 hemoglobin

Chain D:



- Molecule 4: B1 globin chain of giant V2 hemoglobin

Chain H:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	108.87Å 108.87Å 195.16Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.45 – 2.90 42.45 – 2.90	Depositor EDS
% Data completeness (in resolution range)	96.5 (42.45-2.90) 96.4 (42.45-2.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.67 (at 2.90Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.246 , 0.292 0.242 , 0.289	Depositor DCC
$R_{free}$ test set	1437 reflections (5.41%)	DCC
Wilson B-factor (Å <sup>2</sup> )	61.1	Xtriage
Anisotropy	0.165	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 23.8	EDS
Estimated twinning fraction	0.065 for h,-h-k,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 28021 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	9626	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: FUC, HEM, CA, NAG, MAN

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	0.44	0/1178	0.60	0/1585
1	E	0.44	0/1178	0.60	0/1585
2	B	0.40	0/1143	0.59	0/1543
2	F	0.40	0/1143	0.59	0/1543
3	C	0.41	0/1181	0.63	0/1597
3	G	0.41	0/1181	0.63	0/1597
4	D	0.39	0/1164	0.60	0/1579
4	H	0.40	0/1164	0.60	0/1579
All	All	0.41	0/9332	0.60	0/12608

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	1151	0	1123	50	0
1	E	1151	0	1123	50	0
2	B	1120	0	1077	75	0
2	F	1120	0	1077	76	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	1158	0	1115	81	0
3	G	1158	0	1115	83	0
4	D	1140	0	1140	59	0
4	H	1140	0	1140	61	0
5	A	43	0	30	1	0
5	B	43	0	30	1	0
5	C	43	0	30	1	0
5	D	43	0	30	3	0
5	E	43	0	30	1	0
5	F	43	0	30	1	0
5	G	43	0	30	1	0
5	H	43	0	30	3	0
6	B	1	0	0	0	0
6	F	1	0	0	0	0
7	D	71	0	61	0	0
7	H	71	0	61	0	0
All	All	9626	0	9272	461	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 24.

The worst 5 of 461 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:45:ASP:HA	2:B:48:LYS:HE2	1.47	0.95
2:F:45:ASP:HA	2:F:48:LYS:HE2	1.47	0.93
1:E:86:ASP:HB2	4:H:66:ALA:HB1	1.59	0.84
3:C:77:SER:O	3:C:81:LEU:HD23	1.78	0.83
1:A:86:ASP:HB2	4:D:66:ALA:HB1	1.59	0.83

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	144/146 (99%)	126 (88%)	17 (12%)	1 (1%)	30	72
1	E	144/146 (99%)	126 (88%)	17 (12%)	1 (1%)	30	72
2	B	142/144 (99%)	113 (80%)	29 (20%)	0	100	100
2	F	142/144 (99%)	113 (80%)	29 (20%)	0	100	100
3	C	148/150 (99%)	125 (84%)	22 (15%)	1 (1%)	30	72
3	G	148/150 (99%)	125 (84%)	22 (15%)	1 (1%)	30	72
4	D	147/149 (99%)	130 (88%)	14 (10%)	3 (2%)	11	40
4	H	147/149 (99%)	130 (88%)	14 (10%)	3 (2%)	11	40
All	All	1162/1178 (99%)	988 (85%)	164 (14%)	10 (1%)	25	66

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	77	PHE
1	A	77	PHE
4	D	26	ALA
4	D	101	ALA
4	H	26	ALA

### 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	120/120 (100%)	115 (96%)	5 (4%)	40	79
1	E	120/120 (100%)	115 (96%)	5 (4%)	40	79
2	B	116/116 (100%)	110 (95%)	6 (5%)	32	71
2	F	116/116 (100%)	110 (95%)	6 (5%)	32	71
3	C	124/124 (100%)	114 (92%)	10 (8%)	17	45
3	G	124/124 (100%)	114 (92%)	10 (8%)	17	45
4	D	122/122 (100%)	116 (95%)	6 (5%)	35	73
4	H	122/122 (100%)	116 (95%)	6 (5%)	35	73
All	All	964/964 (100%)	910 (94%)	54 (6%)	30	66

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	111	ASP
1	E	70	LEU
4	H	70	ARG
4	D	137	LEU
1	E	4	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	18	GLN
2	F	23	ASN
3	G	34	GLN
3	C	100	GLN
3	C	115	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

Mogul failed to run properly - this section will therefore be empty.

## 5.5 Carbohydrates ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.6 Ligand geometry ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.7 Other polymers ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 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 ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	146/146 (100%)	-0.23	0 100 100	34, 55, 83, 136	0
1	E	146/146 (100%)	-0.33	0 100 100	34, 55, 83, 136	0
2	B	144/144 (100%)	0.07	6 (4%) 35 41	40, 71, 107, 164	0
2	F	144/144 (100%)	-0.11	2 (1%) 72 80	40, 71, 107, 164	0
3	C	150/150 (100%)	0.12	5 (3%) 44 53	39, 69, 113, 170	0
3	G	150/150 (100%)	-0.01	4 (2%) 52 61	39, 69, 113, 170	0
4	D	149/149 (100%)	-0.08	2 (1%) 74 82	40, 66, 104, 140	0
4	H	149/149 (100%)	-0.15	3 (2%) 62 71	40, 66, 104, 140	0
All	All	1178/1178 (100%)	-0.09	22 (1%) 64 72	34, 65, 107, 170	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	G	1	SER	9.3
3	C	1	SER	9.2
4	D	1	SER	5.7
4	H	1	SER	4.7
2	B	137	ILE	3.4

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains.

The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
7	MAN	D	206	11/12	0.83	19.14	218,220,221,222	0
7	NAG	H	202	14/15	0.42	9.57	156,156,156,156	0
7	FUC	H	207	10/11	0.33	7.66	106,106,106,106	0
7	NAG	D	202	14/15	0.29	5.05	156,156,156,156	0
7	FUC	D	207	10/11	0.30	4.76	106,106,106,106	0
7	NAG	D	203	14/15	0.35	3.85	153,153,153,153	0
7	NAG	H	203	14/15	0.43	1.99	153,153,153,153	0
7	MAN	D	205	11/12	0.89	-	236,237,238,239	0
7	MAN	H	206	11/12	0.60	-	218,220,221,222	0
7	MAN	D	204	11/12	0.47	-	230,233,236,237	0
7	MAN	H	205	11/12	0.88	-	236,237,238,239	0
7	MAN	H	204	11/12	0.69	-	230,233,236,237	0

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	HEM	H	201	43/43	0.23	0.32	67,67,67,67	0
6	CA	B	202	1/1	0.24	0.30	67,67,67,67	0
5	HEM	G	200	43/43	0.23	0.26	67,67,67,67	0
5	HEM	F	201	43/43	0.22	0.23	67,67,67,67	0
5	HEM	C	200	43/43	0.23	0.22	67,67,67,67	0
5	HEM	A	200	43/43	0.21	0.15	67,67,67,67	0
5	HEM	D	201	43/43	0.22	-0.00	67,67,67,67	0
5	HEM	E	200	43/43	0.18	-0.10	67,67,67,67	0
5	HEM	B	201	43/43	0.21	-0.33	67,67,67,67	0
6	CA	F	202	1/1	0.04	-6.07	67,67,67,67	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.