



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 04:30 pm BST

PDB ID : 4U1G  
Title : Plasmodium falciparum reticulocyte-binding protein homologue 5 (PfRH5)  
bound to monoclonal antibody QA1  
Authors : Wright, K.E.; Hjerrild, K.A.; Bartlett, J.; Douglas, A.D.; Jin, J.; Brown, R.E.;  
Ashfield, R.; Clemmensen, S.B.; de Jongh, W.A.; Draper, S.J.; Higgins, M.K.  
Deposited on : 2014-07-15  
Resolution : 3.10 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

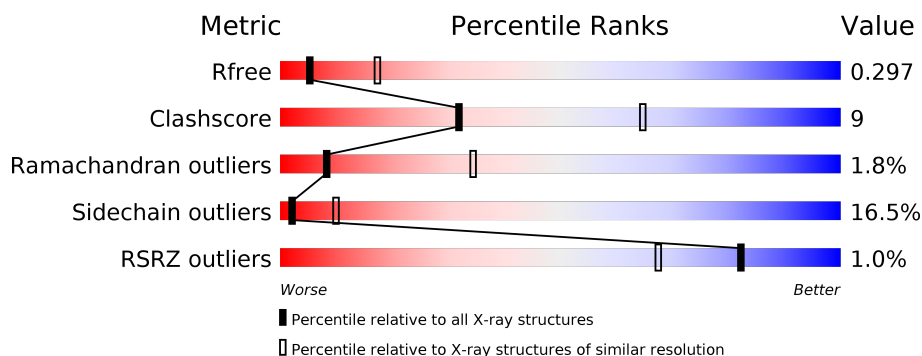
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
$R_{free}$	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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.

Mol	Chain	Length	Quality of chain
1	A	526	<div> <div>36%</div> <div>16%</div> <div>•</div> <div>45%</div> </div>
1	D	526	<div> <div>38%</div> <div>14%</div> <div>•</div> <div>45%</div> </div>
2	B	258	<div> <div>57%</div> <div>21%</div> <div>5%</div> <div>17%</div> </div>
2	E	258	<div> <div>55%</div> <div>23%</div> <div>•</div> <div>17%</div> </div>
3	C	238	<div> <div>63%</div> <div>24%</div> <div>•</div> <div>10%</div> </div>
3	F	238	<div> <div>60%</div> <div>28%</div> <div>•</div> <div>10%</div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 11478 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 Reticulocyte binding protein 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	288	Total	C	N	O	S	0	0	0
			2446	1579	411	441	15			
1	D	289	Total	C	N	O	S	0	0	0
			2458	1588	412	443	15			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	216	ALA	THR	engineered mutation	UNP B2L3N7
D	216	ALA	THR	engineered mutation	UNP B2L3N7

- Molecule 2 is a protein called QA1 monoclonal antibody heavy chain.

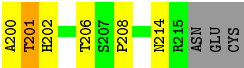
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	213	Total	C	N	O	S	0	0	0
			1611	1020	262	320	9			
2	E	213	Total	C	N	O	S	0	0	0
			1611	1020	262	320	9			

- Molecule 3 is a protein called QA1 monoclonal antibody light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	215	Total	C	N	O	S	0	0	0
			1676	1049	283	339	5			
3	F	215	Total	C	N	O	S	0	0	0
			1676	1049	283	339	5			







● Molecule 3: QA1 monoclonal antibody light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	65.12Å 137.30Å 228.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.57 – 3.10 46.57 – 3.10	Depositor EDS
% Data completeness (in resolution range)	95.1 (46.57-3.10) 95.1 (46.57-3.10)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.16 (at 3.12Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.236 , 0.281 0.253 , 0.297	Depositor DCC
$R_{free}$ test set	1793 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	90.1	Xtriage
Anisotropy	0.677	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 84.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	11478	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.66% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.50	0/2497	0.74	0/3347
1	D	0.51	0/2510	0.75	0/3365
2	B	0.54	0/1651	0.88	1/2248 (0.0%)
2	E	0.57	0/1651	0.89	0/2248
3	C	0.53	0/1719	0.81	0/2340
3	F	0.52	0/1719	0.78	0/2340
All	All	0.53	0/11747	0.80	1/15888 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	159	VAL	N-CA-CB	6.12	124.97	111.50

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	2446	0	2474	53	0
1	D	2458	0	2483	43	0
2	B	1611	0	1559	40	0
2	E	1611	0	1559	33	0
3	C	1676	0	1597	28	0
3	F	1676	0	1597	25	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	11478	0	11269	203	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 9.

The worst 5 of 203 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:B:75:ALA:HB1	2:E:75:ALA:HB1	1.51	0.92
1:D:179:TYR:HD2	1:D:471:THR:HG22	1.35	0.91
1:A:179:TYR:HD2	1:A:471:THR:HG22	1.34	0.90
2:B:11:LEU:HD11	2:B:156:PRO:HG3	1.53	0.90
2:B:153:GLY:HA2	2:B:183:LEU:HB3	1.60	0.81

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 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	284/526 (54%)	261 (92%)	19 (7%)	4 (1%)	11	40
1	D	285/526 (54%)	261 (92%)	22 (8%)	2 (1%)	22	57
2	B	209/258 (81%)	186 (89%)	20 (10%)	3 (1%)	11	40
2	E	209/258 (81%)	182 (87%)	23 (11%)	4 (2%)	8	33
3	C	213/238 (90%)	190 (89%)	15 (7%)	8 (4%)	3	19
3	F	213/238 (90%)	187 (88%)	21 (10%)	5 (2%)	6	28
All	All	1413/2044 (69%)	1267 (90%)	120 (8%)	26 (2%)	8	34

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	200	TYR
3	C	172	SER
2	E	55	ASP
2	E	197	TRP
3	F	55	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	279/511 (55%)	234 (84%)	45 (16%)	2	10
1	D	280/511 (55%)	237 (85%)	43 (15%)	2	11
2	B	182/219 (83%)	146 (80%)	36 (20%)	1	5
2	E	182/219 (83%)	151 (83%)	31 (17%)	2	9
3	C	190/211 (90%)	162 (85%)	28 (15%)	3	13
3	F	190/211 (90%)	158 (83%)	32 (17%)	2	9
All	All	1303/1882 (69%)	1088 (84%)	215 (16%)	2	10

5 of 215 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	146	LYS
1	D	226	ASP
3	F	108	LEU
3	C	159	ARG
1	D	162	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	141	ASN
3	C	142	ASN
3	F	80	HIS
3	C	94	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	C	128	GLN

### 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](#)

There are no ligands in this entry.

### 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](#)

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	288/526 (54%)	0.02	7 (2%) 59 37	67, 108, 171, 199	0
1	D	289/526 (54%)	0.12	6 (2%) 63 43	77, 113, 174, 203	0
2	B	213/258 (82%)	-0.14	1 (0%) 91 81	51, 84, 123, 136	0
2	E	213/258 (82%)	-0.17	0 100 100	51, 84, 119, 142	0
3	C	215/238 (90%)	-0.11	0 100 100	57, 90, 122, 141	1 (0%)
3	F	215/238 (90%)	0.06	0 100 100	55, 99, 129, 143	1 (0%)
All	All	1433/2044 (70%)	-0.03	14 (0%) 82 67	51, 97, 153, 203	2 (0%)

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	404	TYR	3.7
1	D	160	SER	3.3
1	D	325	PHE	3.1
1	A	226	ASP	3.1
1	D	171	LEU	2.9

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.