



# Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 06:31 am BST

PDB ID : 4UV6  
Title : Crystal structure of apical membrane antigen 1 from Plasmodium knowlesi  
Authors : Vulliez-Le Normand, B.; Saul, F.A.; Bentley, G.A.  
Deposited on : 2014-08-04  
Resolution : 2.45 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

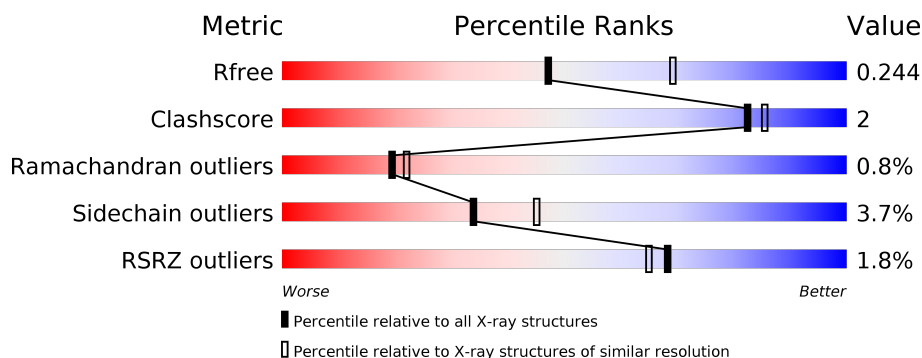
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.45 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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	370	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: -5px; left: 0; width: 100%; height: 100%; background: linear-gradient(to right, red 0%, orange 1%, yellow 82%, green 99%, grey 100%);"></div> <div style="position: absolute; top: 5px; left: 0; width: 100%; height: 100%; background: linear-gradient(to right, red 0%, orange 1%, yellow 82%, green 99%, grey 100%);"></div> </div> <div> <span style="display: inline-block; width: 100%; text-align: center;">82% 9% 9%</span> </div> </div>
1	B	370	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: -5px; left: 0; width: 100%; height: 100%; background: linear-gradient(to right, red 0%, orange 2%, yellow 80%, green 91%, grey 100%);"></div> <div style="position: absolute; top: 5px; left: 0; width: 100%; height: 100%; background: linear-gradient(to right, red 0%, orange 2%, yellow 80%, green 91%, grey 100%);"></div> </div> <div> <span style="display: inline-block; width: 100%; text-align: center;">80% 11% 9%</span> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5809 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 APICAL MEROZOITE ANTIGEN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	336	Total	C	N	O	S	0	0	0
			2704	1700	467	521	16			
1	B	336	Total	C	N	O	S	0	2	0
			2719	1709	473	522	15			

There are 58 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	41	GLU	-	cloning artifact	UNP B3L5E1
A	42	PHE	-	cloning artifact	UNP B3L5E1
A	388	GLY	-	cloning artifact	UNP B3L5E1
A	389	LEU	-	expression tag	UNP B3L5E1
A	390	GLU	-	expression tag	UNP B3L5E1
A	391	GLN	-	expression tag	UNP B3L5E1
A	392	LYS	-	expression tag	UNP B3L5E1
A	393	LEU	-	expression tag	UNP B3L5E1
A	394	ILE	-	expression tag	UNP B3L5E1
A	395	SER	-	expression tag	UNP B3L5E1
A	396	GLU	-	expression tag	UNP B3L5E1
A	397	GLU	-	expression tag	UNP B3L5E1
A	398	ASP	-	expression tag	UNP B3L5E1
A	399	LEU	-	expression tag	UNP B3L5E1
A	400	ASN	-	expression tag	UNP B3L5E1
A	401	SER	-	expression tag	UNP B3L5E1
A	402	ALA	-	expression tag	UNP B3L5E1
A	403	VAL	-	expression tag	UNP B3L5E1
A	404	ASP	-	expression tag	UNP B3L5E1
A	405	HIS	-	expression tag	UNP B3L5E1
A	406	HIS	-	expression tag	UNP B3L5E1
A	407	HIS	-	expression tag	UNP B3L5E1
A	408	HIS	-	expression tag	UNP B3L5E1
A	409	HIS	-	expression tag	UNP B3L5E1
A	410	HIS	-	expression tag	UNP B3L5E1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	107	LYS	ASN	engineered mutation	UNP B3L5E1
A	178	ASN	SER	engineered mutation	UNP B3L5E1
A	189	GLU	ASN	engineered mutation	UNP B3L5E1
A	240	ARG	SER	engineered mutation	UNP B3L5E1
B	41	GLU	-	cloning artifact	UNP B3L5E1
B	42	PHE	-	cloning artifact	UNP B3L5E1
B	388	GLY	-	cloning artifact	UNP B3L5E1
B	389	LEU	-	expression tag	UNP B3L5E1
B	390	GLU	-	expression tag	UNP B3L5E1
B	391	GLN	-	expression tag	UNP B3L5E1
B	392	LYS	-	expression tag	UNP B3L5E1
B	393	LEU	-	expression tag	UNP B3L5E1
B	394	ILE	-	expression tag	UNP B3L5E1
B	395	SER	-	expression tag	UNP B3L5E1
B	396	GLU	-	expression tag	UNP B3L5E1
B	397	GLU	-	expression tag	UNP B3L5E1
B	398	ASP	-	expression tag	UNP B3L5E1
B	399	LEU	-	expression tag	UNP B3L5E1
B	400	ASN	-	expression tag	UNP B3L5E1
B	401	SER	-	expression tag	UNP B3L5E1
B	402	ALA	-	expression tag	UNP B3L5E1
B	403	VAL	-	expression tag	UNP B3L5E1
B	404	ASP	-	expression tag	UNP B3L5E1
B	405	HIS	-	expression tag	UNP B3L5E1
B	406	HIS	-	expression tag	UNP B3L5E1
B	407	HIS	-	expression tag	UNP B3L5E1
B	408	HIS	-	expression tag	UNP B3L5E1
B	409	HIS	-	expression tag	UNP B3L5E1
B	410	HIS	-	expression tag	UNP B3L5E1
B	107	LYS	ASN	engineered mutation	UNP B3L5E1
B	178	ASN	SER	engineered mutation	UNP B3L5E1
B	189	GLU	ASN	engineered mutation	UNP B3L5E1
B	240	ARG	SER	engineered mutation	UNP B3L5E1

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	213	Total O 213 213	0	0
2	B	173	Total O 173 173	0	0



• Molecule 1: APICAL MEROZOITE ANTIGEN 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.32Å 105.70Å 104.75Å 90.00° 98.04° 90.00°	Depositor
Resolution (Å)	47.09 – 2.45 47.09 – 2.45	Depositor EDS
% Data completeness (in resolution range)	97.1 (47.09-2.45) 97.1 (47.09-2.45)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 2.45Å)	Xtriage
Refinement program	BUSTER 2.11.4	Depositor
R, $R_{free}$	0.171 , 0.227 0.188 , 0.244	Depositor DCC
$R_{free}$ test set	843 reflections (2.42%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.3	Xtriage
Anisotropy	0.561	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 69.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5809	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.15% 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

### 5.1 Standard geometry

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	0.53	0/2763	0.71	0/3726
1	B	0.50	0/2785	0.70	0/3757
All	All	0.51	0/5548	0.71	0/7483

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

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	2704	0	2582	13	0
1	B	2719	0	2601	13	0
2	A	213	0	0	0	0
2	B	173	0	0	1	0
All	All	5809	0	5183	26	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 2.

All (26) 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:A:288:GLU:HG3	1:A:296:ARG:HH22	1.62	0.65

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:78:GLU:HB2	1:A:91:ALA:HB2	1.78	0.64
1:A:201:GLU:HB2	1:A:223:PRO:HD3	1.80	0.63
1:B:78:GLU:HB2	1:B:91:ALA:HB2	1.80	0.63
1:B:322:SER:HA	1:B:325:LEU:HD12	1.82	0.61
1:A:322:SER:HA	1:A:325:LEU:HD12	1.82	0.61
1:B:201:GLU:HB2	1:B:223:PRO:HD3	1.83	0.60
1:A:55:TRP:CD2	1:A:389:LEU:HD11	2.46	0.50
1:B:182:PRO:HG2	1:B:195:LEU:HB2	1.94	0.49
1:B:55:TRP:CD2	1:B:389:LEU:HD11	2.48	0.47
1:A:182:PRO:HG2	1:A:195:LEU:HB2	1.95	0.47
1:A:113:PRO:HB3	1:A:129:PRO:HA	1.97	0.46
1:A:257:PHE:HB3	1:A:361:PRO:HB3	1.98	0.46
1:B:257:PHE:HB3	1:B:361:PRO:HB3	1.99	0.44
1:B:300:GLU:HB2	1:B:303:THR:HG23	1.99	0.44
1:B:252:LEU:HB2	1:B:372:ALA:HB3	2.00	0.43
1:A:288:GLU:HG3	1:A:296:ARG:NH2	2.30	0.43
1:A:159:ILE:HG23	1:A:235:LEU:HD21	2.01	0.43
1:A:310:GLU:OE1	1:A:313:ARG:HD3	2.19	0.42
1:B:64:LEU:HD12	1:B:343:ASN:HB3	2.03	0.41
1:A:152:ASP:O	1:A:155:LYS:HG2	2.21	0.41
1:B:93:LYS:HB3	1:B:240:ARG:HG3	2.03	0.41
1:B:159:ILE:HG23	1:B:235:LEU:HD21	2.01	0.41
1:B:134:HIS:HB3	2:B:2055:HOH:O	2.19	0.41
1:B:205:PRO:HD3	1:B:297:GLN:HB3	2.01	0.41
1:A:152:ASP:HA	1:A:155:LYS:HE3	2.03	0.40

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	328/370 (89%)	311 (95%)	15 (5%)	2 (1%)	25	29

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	332/370 (90%)	314 (95%)	15 (4%)	3 (1%)	17	19
All	All	660/740 (89%)	625 (95%)	30 (4%)	5 (1%)	19	22

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	176	ASN
1	B	176	ASN
1	A	210	PRO
1	B	134	HIS
1	B	210	PRO

### 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	294/325 (90%)	285 (97%)	9 (3%)	40	52
1	B	295/325 (91%)	282 (96%)	13 (4%)	28	37
All	All	589/650 (91%)	567 (96%)	22 (4%)	34	45

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	128	PHE
1	A	163	LYS
1	A	170	VAL
1	A	173	GLU
1	A	196	TYR
1	A	203	MET
1	A	236	SER
1	A	277	ARG
1	A	314	GLN
1	B	107	LYS
1	B	128	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	132	ASP
1	B	133	ASP
1	B	163	LYS
1	B	170	VAL
1	B	175	GLN
1	B	196	TYR
1	B	236	SER
1	B	277	ARG
1	B	314	GLN
1	B	395	SER
1	B	398	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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	336/370 (90%)	-0.35	4 (1%) 79 77	18, 36, 76, 106	0
1	B	336/370 (90%)	-0.22	8 (2%) 59 54	22, 43, 90, 118	0
All	All	672/740 (90%)	-0.28	12 (1%) 68 65	18, 40, 82, 118	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	118	ALA	4.8
1	B	211	ASP	4.5
1	B	207	TYR	4.3
1	A	175	GLN	3.1
1	B	305	TYR	3.1
1	B	302	LEU	2.9
1	B	107	LYS	2.6
1	A	118	ALA	2.6
1	B	297	GLN	2.2
1	A	176	ASN	2.2
1	A	211	ASP	2.2
1	B	120	ARG	2.0

### 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.