



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2024 – 02:55 pm BST

PDB ID : 9EVO  
Title : Plasmodium falciparum apical membrane antigen 3D7  
Authors : Bentley, G.A.; Saul, F.A.; Vulliez-Le Normand, B.  
Deposited on : 2024-03-31  
Resolution : 2.10 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

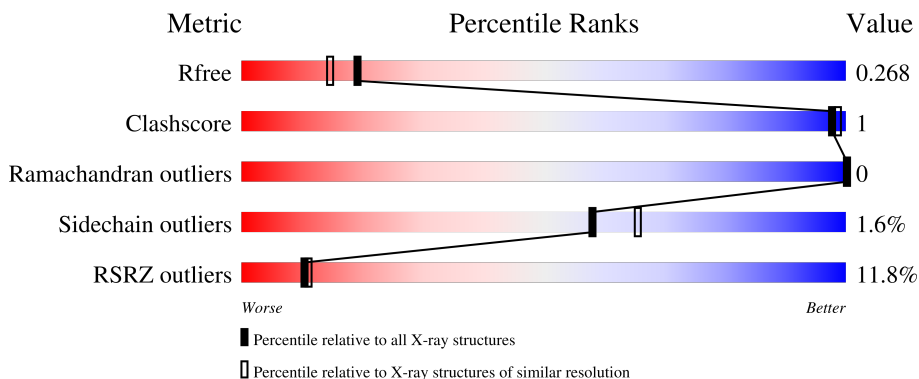
# 1 Overall quality at a glance

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))
$R_{free}$	164625	6234 (2.10-2.10)
Clashscore	180529	6893 (2.10-2.10)
Ramachandran outliers	177936	6839 (2.10-2.10)
Sidechain outliers	177891	6840 (2.10-2.10)
RSRZ outliers	164620	6234 (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 of 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	363	 9% 87% 11%
1	B	363	 10% 76% 21%

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 5102 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 membrane antigen 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	323	2524	1602	424	482	16	0	0	0
1	B	286	2256	1435	373	433	15	0	1	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	95	MET	-	initiating methionine	UNP Q7KQK5
A	96	ALA	-	expression tag	UNP Q7KQK5
A	164	ALA	THR	engineered mutation	UNP Q7KQK5
A	288	ALA	THR	engineered mutation	UNP Q7KQK5
A	373	ALA	SER	engineered mutation	UNP Q7KQK5
A	423	ALA	SER	engineered mutation	UNP Q7KQK5
A	424	ALA	SER	conflict	UNP Q7KQK5
A	443	ALA	-	expression tag	UNP Q7KQK5
A	444	ALA	-	expression tag	UNP Q7KQK5
A	445	ALA	-	expression tag	UNP Q7KQK5
A	446	HIS	-	expression tag	UNP Q7KQK5
A	447	HIS	-	expression tag	UNP Q7KQK5
A	448	HIS	-	expression tag	UNP Q7KQK5
A	449	HIS	-	expression tag	UNP Q7KQK5
A	450	HIS	-	expression tag	UNP Q7KQK5
A	451	HIS	-	expression tag	UNP Q7KQK5
A	452	SER	-	expression tag	UNP Q7KQK5
A	453	GLU	-	expression tag	UNP Q7KQK5
A	454	LYS	-	expression tag	UNP Q7KQK5
A	455	ASP	-	expression tag	UNP Q7KQK5
A	456	GLU	-	expression tag	UNP Q7KQK5
A	457	LEU	-	expression tag	UNP Q7KQK5
B	95	MET	-	initiating methionine	UNP Q7KQK5
B	96	ALA	-	expression tag	UNP Q7KQK5
B	164	ALA	THR	engineered mutation	UNP Q7KQK5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	288	ALA	THR	engineered mutation	UNP Q7KQK5
B	373	ALA	SER	engineered mutation	UNP Q7KQK5
B	423	ALA	SER	engineered mutation	UNP Q7KQK5
B	424	ALA	SER	conflict	UNP Q7KQK5
B	443	ALA	-	expression tag	UNP Q7KQK5
B	444	ALA	-	expression tag	UNP Q7KQK5
B	445	ALA	-	expression tag	UNP Q7KQK5
B	446	HIS	-	expression tag	UNP Q7KQK5
B	447	HIS	-	expression tag	UNP Q7KQK5
B	448	HIS	-	expression tag	UNP Q7KQK5
B	449	HIS	-	expression tag	UNP Q7KQK5
B	450	HIS	-	expression tag	UNP Q7KQK5
B	451	HIS	-	expression tag	UNP Q7KQK5
B	452	SER	-	expression tag	UNP Q7KQK5
B	453	GLU	-	expression tag	UNP Q7KQK5
B	454	LYS	-	expression tag	UNP Q7KQK5
B	455	ASP	-	expression tag	UNP Q7KQK5
B	456	GLU	-	expression tag	UNP Q7KQK5
B	457	LEU	-	expression tag	UNP Q7KQK5

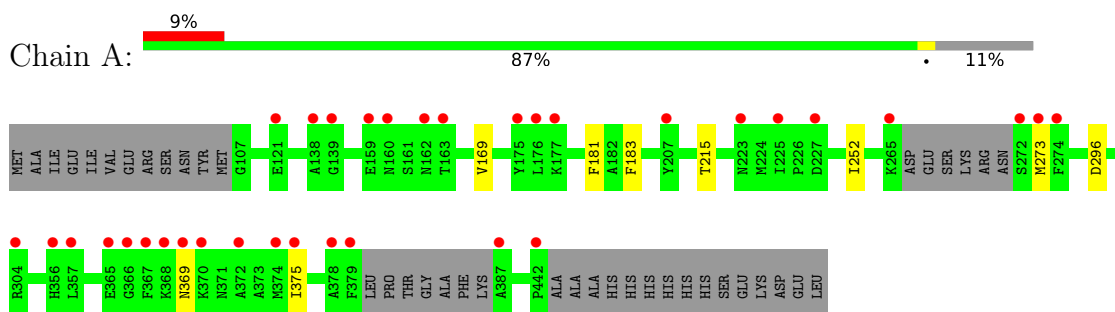
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	158	Total O 158 158	0	0
2	B	164	Total O 164 164	0	0

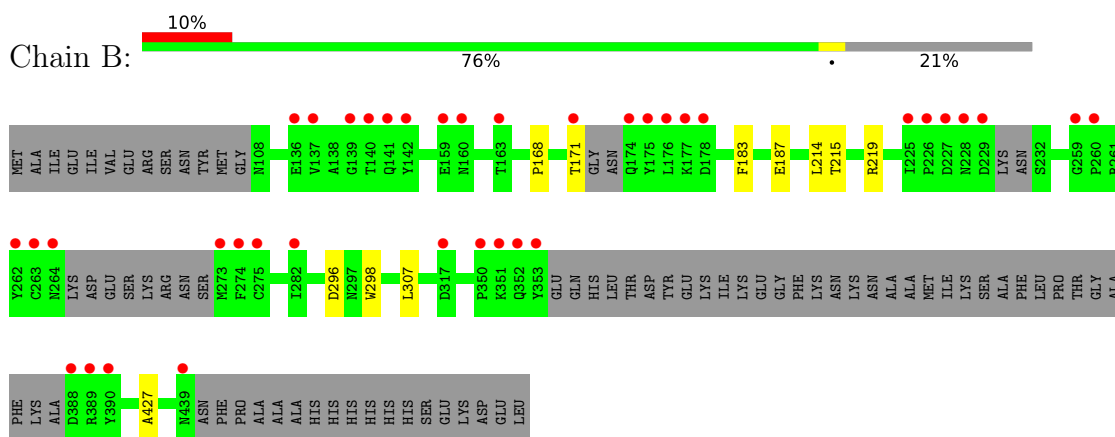
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 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: Apical membrane antigen 1



- Molecule 1: Apical membrane antigen 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	38.23Å 62.06Å 71.50Å 89.13° 89.90° 83.92°	Depositor
Resolution (Å)	46.37 – 2.10 46.37 – 2.10	Depositor EDS
% Data completeness (in resolution range)	97.8 (46.37-2.10) 97.8 (46.37-2.10)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.13 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.8.0415	Depositor
R, $R_{free}$	0.207 , 0.256 0.216 , 0.268	Depositor DCC
$R_{free}$ test set	1908 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.9	Xtrriage
Anisotropy	0.034	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 44.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.025 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	5102	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.34% 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	0.46	0/2588	0.82	0/3511
1	B	0.47	0/2319	0.81	0/3147
All	All	0.47	0/4907	0.81	0/6658

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	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	219	ARG	Sidechain

### 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	2524	0	2363	3	0
1	B	2256	0	2109	3	0
2	A	158	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	164	0	0	0	0
All	All	5102	0	4472	6	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 1.

All (6) 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:B:215:THR:HG21	1:B:296:ASP:HA	1.91	0.52
1:B:307:LEU:HB2	1:B:427:ALA:HB3	1.96	0.48
1:A:181:PHE:HA	1:A:252:ILE:HD12	1.98	0.44
1:A:215:THR:HG21	1:A:296:ASP:HA	2.02	0.42
1:A:252:ILE:HG23	1:A:375:ILE:HG23	2.02	0.42
1:B:214:LEU:HD12	1:B:298:TRP:CZ2	2.56	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	317/363 (87%)	308 (97%)	9 (3%)	0	100	100
1	B	277/363 (76%)	266 (96%)	11 (4%)	0	100	100
All	All	594/726 (82%)	574 (97%)	20 (3%)	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 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	265/315 (84%)	261 (98%)	4 (2%)	60	67
1	B	242/315 (77%)	238 (98%)	4 (2%)	56	63
All	All	507/630 (80%)	499 (98%)	8 (2%)	58	65

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	169	VAL
1	A	183	PHE
1	A	273	MET
1	A	369	ASN
1	B	168	PRO
1	B	171	THR
1	B	183	PHE
1	B	187	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	160	ASN
1	A	233	ASN
1	A	257	ASN
1	A	258	ASN
1	A	286	ASN
1	A	327	ASN
1	B	160	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 oligosaccharides 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

### 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	323/363 (88%)	0.68	34 (10%) <b>13</b> <b>14</b>	11, 25, 55, 77	1 (0%)
1	B	286/363 (78%)	0.79	38 (13%) <b>8</b> <b>9</b>	15, 25, 69, 96	1 (0%)
All	All	609/726 (83%)	0.73	72 (11%) <b>10</b> <b>11</b>	11, 25, 61, 96	2 (0%)

All (72) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	366	GLY	6.1
1	B	274	PHE	5.8
1	A	375	ILE	5.6
1	A	379	PHE	5.6
1	B	177	LYS	5.2
1	B	352	GLN	4.8
1	B	264	ASN	4.7
1	B	227	ASP	4.7
1	B	273	MET	4.5
1	B	262	TYR	4.5
1	B	260	PRO	4.4
1	B	353	TYR	4.3
1	A	369	ASN	4.3
1	A	387	ALA	4.1
1	B	175	TYR	4.1
1	B	228	ASN	4.1
1	A	357	LEU	4.0
1	B	390	TYR	4.0
1	A	162	ASN	3.7
1	A	367	PHE	3.7
1	A	274	PHE	3.6
1	B	388	ASP	3.6
1	A	272	SER	3.5
1	B	171	THR	3.4

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	351	LYS	3.4
1	B	229	ASP	3.4
1	A	225	ILE	3.3
1	A	176	LEU	3.3
1	B	176	LEU	3.3
1	B	160	ASN	3.2
1	B	178	ASP	3.2
1	A	378	ALA	3.2
1	A	160	ASN	3.2
1	A	139	GLY	3.1
1	A	175	TYR	3.1
1	A	207	TYR	3.1
1	B	263	CYS	3.1
1	B	163	THR	3.0
1	B	174	GLN	2.9
1	A	273	MET	2.8
1	B	389	ARG	2.8
1	B	140	THR	2.7
1	B	350	PRO	2.7
1	A	365	GLU	2.7
1	B	439	ASN	2.6
1	B	159	GLU	2.6
1	A	368	LYS	2.6
1	A	442	PRO	2.6
1	A	356	HIS	2.6
1	A	163	THR	2.5
1	B	136	GLU	2.5
1	A	265	LYS	2.4
1	B	139	GLY	2.4
1	A	121	GLU	2.3
1	A	372	ALA	2.3
1	B	226	PRO	2.3
1	A	370	LYS	2.3
1	B	275	CYS	2.3
1	B	137	VAL	2.3
1	B	141	GLN	2.3
1	A	374	MET	2.3
1	B	225	ILE	2.2
1	A	304	ARG	2.2
1	B	142	TYR	2.2
1	A	177	LYS	2.2
1	B	259	GLY	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	223	ASN	2.2
1	A	138	ALA	2.1
1	B	317	ASP	2.1
1	B	282	ILE	2.1
1	A	159	GLU	2.0
1	A	227	ASP	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 monosaccharides 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.