



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 23, 2017 – 05:55 PM EST

PDB ID : 5TIK
Title : Structural basis for inhibition of erythrocyte invasion by antibodies to Plasmodium falciparum protein CyRPA
Authors : Chen, L.; Xu, Y.; Wang, W.; Thompson, J.K.; Goddard-Borger, E.; Lawrence, M.C.; Cowman, A.F.
Deposited on : 2016-10-03
Resolution : 3.09 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

A user guide is available at

<http://wwpdb.org/validation/2016/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
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

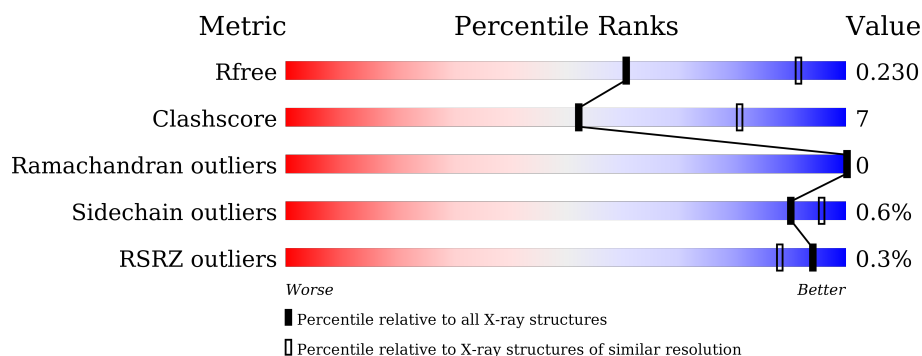
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.09 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}	91344	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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 ≥ 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 $\leq 5\%$. 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	335	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 19%, green 74%, grey 7%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 74% 19% 7% </div> </div>
1	B	335	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 77%, yellow 20%, grey 3%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 77% 20% • </div> </div>
1	C	335	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 78%, yellow 17%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 78% 17% • </div> </div>
1	D	335	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 78%, yellow 16%, orange 5%, grey 1%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 78% 16% • 5% </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10627 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 Cysteine-rich protective antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	310	Total	C	N	O	S	0	0	0
			2585	1668	415	489	13			
1	B	324	Total	C	N	O	S	0	0	0
			2698	1736	436	513	13			
1	C	321	Total	C	N	O	S	0	0	0
			2680	1725	429	513	13			
1	D	318	Total	C	N	O	S	0	0	0
			2651	1706	429	503	13			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q8IFM8
A	0	THR	-	expression tag	UNP Q8IFM8
A	118	ALA	SER	conflict	UNP Q8IFM8
A	295	ALA	THR	conflict	UNP Q8IFM8
A	311	ALA	THR	conflict	UNP Q8IFM8
B	-1	GLY	-	expression tag	UNP Q8IFM8
B	0	THR	-	expression tag	UNP Q8IFM8
B	118	ALA	SER	conflict	UNP Q8IFM8
B	295	ALA	THR	conflict	UNP Q8IFM8
B	311	ALA	THR	conflict	UNP Q8IFM8
C	-1	GLY	-	expression tag	UNP Q8IFM8
C	0	THR	-	expression tag	UNP Q8IFM8
C	118	ALA	SER	conflict	UNP Q8IFM8
C	295	ALA	THR	conflict	UNP Q8IFM8
C	311	ALA	THR	conflict	UNP Q8IFM8
D	-1	GLY	-	expression tag	UNP Q8IFM8
D	0	THR	-	expression tag	UNP Q8IFM8
D	118	ALA	SER	conflict	UNP Q8IFM8
D	295	ALA	THR	conflict	UNP Q8IFM8
D	311	ALA	THR	conflict	UNP Q8IFM8

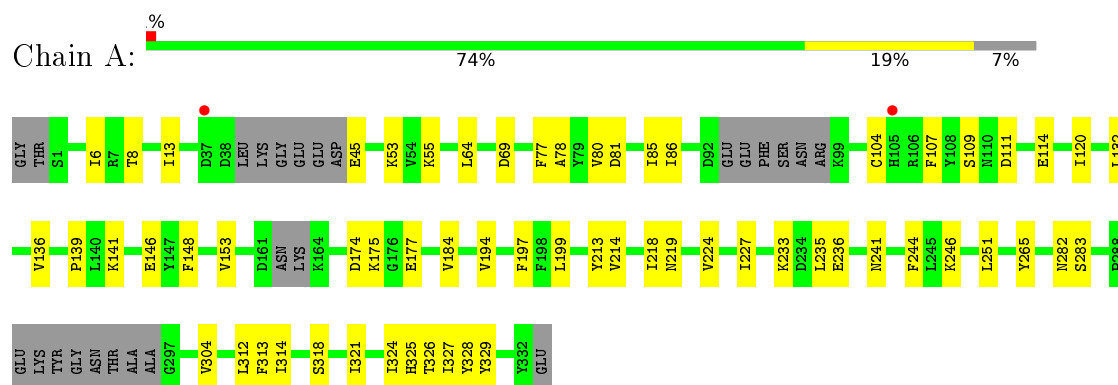
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	5	Total 5	O 5	0	0
2	B	1	Total 1	O 1	0	0
2	C	5	Total 5	O 5	0	0
2	D	2	Total 2	O 2	0	0

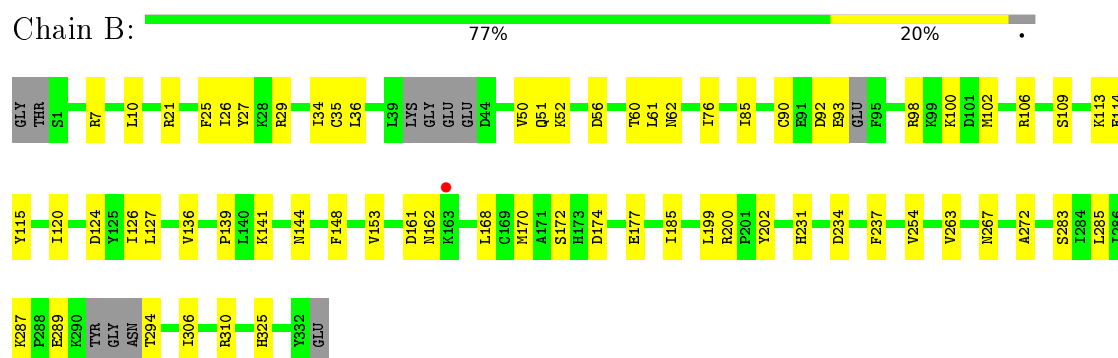
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

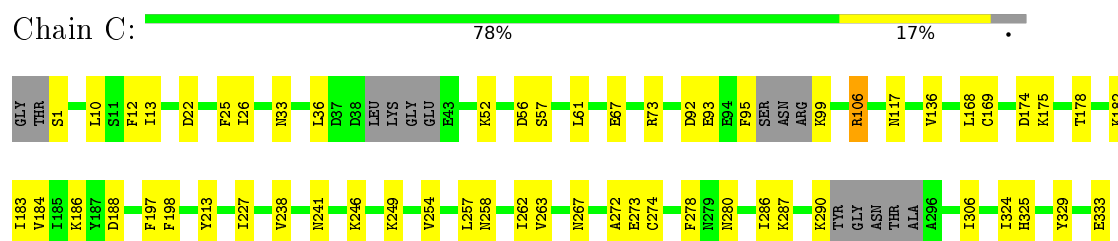
- Molecule 1: Cysteine-rich protective antigen



- Molecule 1: Cysteine-rich protective antigen



- Molecule 1: Cysteine-rich protective antigen



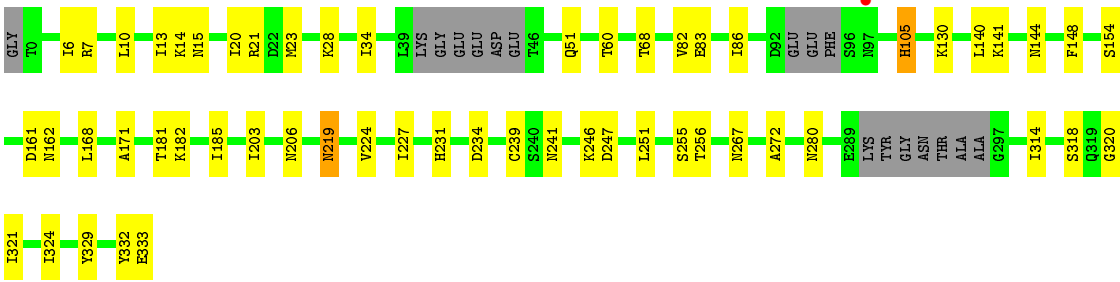
- Molecule 1: Cysteine-rich protective antigen

Chain D:

78%

16%

• 5%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	68.70Å 83.56Å 95.32Å 96.77° 104.11° 115.21°	Depositor
Resolution (Å)	48.26 – 3.09 48.26 – 3.09	Depositor EDS
% Data completeness (in resolution range)	97.8 (48.26-3.09) 89.4 (48.26-3.09)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.01 (at 3.07Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.187 , 0.229 0.186 , 0.230	Depositor DCC
R_{free} test set	1582 reflections (4.91%)	DCC
Wilson B-factor (Å ²)	71.1	Xtriage
Anisotropy	0.525	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 53.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	10627	wwPDB-VP
Average B, all atoms (Å ²)	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.08% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.25	0/2643	0.45	0/3566
1	B	0.25	0/2758	0.44	0/3721
1	C	0.25	0/2740	0.44	0/3696
1	D	0.25	0/2710	0.44	0/3657
All	All	0.25	0/10851	0.44	0/14640

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	2585	0	2491	39	0
1	B	2698	0	2601	38	0
1	C	2680	0	2572	36	0
1	D	2651	0	2559	31	0
2	A	5	0	0	2	0
2	B	1	0	0	0	0
2	C	5	0	0	0	0
2	D	2	0	0	0	0
All	All	10627	0	10223	139	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 7.

The worst 5 of 139 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:93:GLU:HG2	1:B:98:ARG:HD3	1.61	0.81
1:A:78:ALA:HB3	1:A:136:VAL:HG12	1.68	0.76
1:B:7:ARG:HH21	1:C:333:GLU:HG3	1.54	0.73
1:C:1:SER:N	1:D:10:LEU:O	2.23	0.72
1:C:106:ARG:NH1	1:C:175:LYS:O	2.24	0.71

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	300/335 (90%)	284 (95%)	16 (5%)	0	100	100
1	B	316/335 (94%)	303 (96%)	13 (4%)	0	100	100
1	C	313/335 (93%)	297 (95%)	16 (5%)	0	100	100
1	D	310/335 (92%)	295 (95%)	15 (5%)	0	100	100
All	All	1239/1340 (92%)	1179 (95%)	60 (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 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	292/312 (94%)	291 (100%)	1 (0%)	94	97
1	B	304/312 (97%)	302 (99%)	2 (1%)	88	95
1	C	302/312 (97%)	301 (100%)	1 (0%)	94	97
1	D	300/312 (96%)	297 (99%)	3 (1%)	82	93
All	All	1198/1248 (96%)	1191 (99%)	7 (1%)	90	95

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	106	ARG
1	D	247	ASP
1	D	105	HIS
1	B	29	ARG
1	D	219	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	33	ASN
1	C	51	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, 95th 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 ²)	Q<0.9
1	A	310/335 (92%)	-0.28	2 (0%) 90 80	38, 79, 131, 153	0
1	B	324/335 (96%)	-0.31	1 (0%) 94 88	44, 83, 132, 150	0
1	C	321/335 (95%)	-0.39	0 100 100	40, 78, 124, 157	0
1	D	318/335 (94%)	-0.31	1 (0%) 94 88	32, 75, 130, 180	0
All	All	1273/1340 (95%)	-0.33	4 (0%) 94 88	32, 78, 130, 180	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	97	ASN	2.6
1	A	105	HIS	2.6
1	A	37	ASP	2.4
1	B	163	LYS	2.2

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.