



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:28 PM GMT

PDB ID : 3ZLD
Title : Crystal structure of Toxoplasma gondii sporozoite AMA1 in complex with a
36 aa region of sporozoite RON2
Authors : Tonkin, M.L.; Boulanger, M.J.
Deposited on : 2013-01-30
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
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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
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) : trunk26865

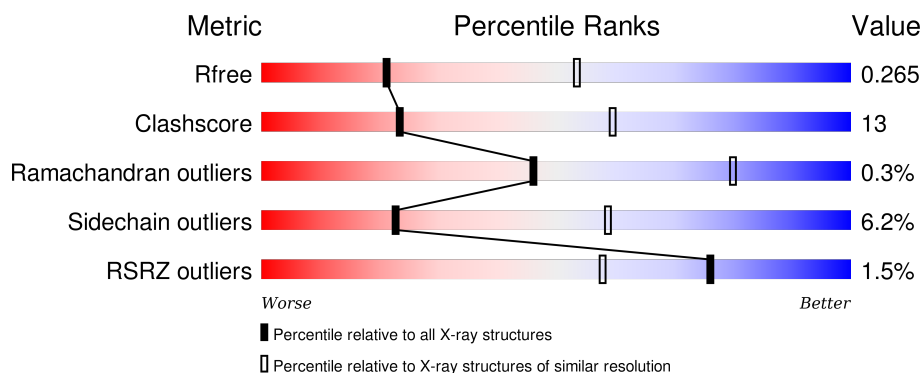
1 Overall quality at a glance

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}	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	396	 2% 68% 22% • 8%
2	B	40	 53% 20% • 25%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2981 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	363	2745	1728	455	544	18	0	0	0

There are 104 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	92	GLY	-	EXPRESSION TAG	UNP B6K9M7
A	93	SER	-	EXPRESSION TAG	UNP B6K9M7
A	94	ALA	-	EXPRESSION TAG	UNP B6K9M7
A	95	MET	-	EXPRESSION TAG	UNP B6K9M7
A	96	GLY	-	EXPRESSION TAG	UNP B6K9M7
A	389	GLY	-	SEE REMARK 999	UNP B6K9M7
A	390	PHE	-	SEE REMARK 999	UNP B6K9M7
A	391	GLY	-	SEE REMARK 999	UNP B6K9M7
A	392	ALA	-	SEE REMARK 999	UNP B6K9M7
A	393	ASN	-	SEE REMARK 999	UNP B6K9M7
A	394	TRP	-	SEE REMARK 999	UNP B6K9M7
A	395	ALA	-	SEE REMARK 999	UNP B6K9M7
A	396	ASN	-	SEE REMARK 999	UNP B6K9M7
A	397	PHE	-	SEE REMARK 999	UNP B6K9M7
A	398	TYR	-	SEE REMARK 999	UNP B6K9M7
A	399	LEU	-	SEE REMARK 999	UNP B6K9M7
A	400	GLU	-	SEE REMARK 999	UNP B6K9M7
A	401	LYS	-	SEE REMARK 999	UNP B6K9M7
A	402	GLU	-	SEE REMARK 999	UNP B6K9M7
A	403	SER	-	SEE REMARK 999	UNP B6K9M7
A	404	GLY	-	SEE REMARK 999	UNP B6K9M7
A	405	GLU	-	SEE REMARK 999	UNP B6K9M7
A	406	THR	-	SEE REMARK 999	UNP B6K9M7
A	407	ILE	-	SEE REMARK 999	UNP B6K9M7
A	408	CYS	-	SEE REMARK 999	UNP B6K9M7
A	409	ALA	-	SEE REMARK 999	UNP B6K9M7
A	410	ILE	-	SEE REMARK 999	UNP B6K9M7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	411	PHE	-	SEE REMARK 999	UNP B6K9M7
A	412	ASP	-	SEE REMARK 999	UNP B6K9M7
A	413	GLN	-	SEE REMARK 999	UNP B6K9M7
A	414	VAL	-	SEE REMARK 999	UNP B6K9M7
A	415	PRO	-	SEE REMARK 999	UNP B6K9M7
A	416	ASP	-	SEE REMARK 999	UNP B6K9M7
A	417	CYS	-	SEE REMARK 999	UNP B6K9M7
A	418	PHE	-	SEE REMARK 999	UNP B6K9M7
A	419	ALA	-	SEE REMARK 999	UNP B6K9M7
A	420	PRO	-	SEE REMARK 999	UNP B6K9M7
A	421	ILE	-	SEE REMARK 999	UNP B6K9M7
A	422	THR	-	SEE REMARK 999	UNP B6K9M7
A	423	GLY	-	SEE REMARK 999	UNP B6K9M7
A	424	ALA	-	SEE REMARK 999	UNP B6K9M7
A	425	VAL	-	SEE REMARK 999	UNP B6K9M7
A	426	ALA	-	SEE REMARK 999	UNP B6K9M7
A	427	TYR	-	SEE REMARK 999	UNP B6K9M7
A	428	THR	-	SEE REMARK 999	UNP B6K9M7
A	429	ALA	-	SEE REMARK 999	UNP B6K9M7
A	430	LEU	-	SEE REMARK 999	UNP B6K9M7
A	431	GLY	-	SEE REMARK 999	UNP B6K9M7
A	432	SER	-	SEE REMARK 999	UNP B6K9M7
A	433	SER	-	SEE REMARK 999	UNP B6K9M7
A	434	THR	-	SEE REMARK 999	UNP B6K9M7
A	435	GLU	-	SEE REMARK 999	UNP B6K9M7
A	436	VAL	-	SEE REMARK 999	UNP B6K9M7
A	437	ASN	-	SEE REMARK 999	UNP B6K9M7
A	438	LEU	-	SEE REMARK 999	UNP B6K9M7
A	439	PRO	-	SEE REMARK 999	UNP B6K9M7
A	440	GLN	-	SEE REMARK 999	UNP B6K9M7
A	441	CYS	-	SEE REMARK 999	UNP B6K9M7
A	442	ASP	-	SEE REMARK 999	UNP B6K9M7
A	443	SER	-	SEE REMARK 999	UNP B6K9M7
A	444	ALA	-	SEE REMARK 999	UNP B6K9M7
A	445	SER	-	SEE REMARK 999	UNP B6K9M7
A	446	PHE	-	SEE REMARK 999	UNP B6K9M7
A	447	ILE	-	SEE REMARK 999	UNP B6K9M7
A	448	PRO	-	SEE REMARK 999	UNP B6K9M7
A	449	ILE	-	SEE REMARK 999	UNP B6K9M7
A	450	GLU	-	SEE REMARK 999	UNP B6K9M7
A	451	GLY	-	SEE REMARK 999	UNP B6K9M7
A	452	PRO	-	SEE REMARK 999	UNP B6K9M7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	453	CYS	-	SEE REMARK 999	UNP B6K9M7
A	454	ASN	-	SEE REMARK 999	UNP B6K9M7
A	455	ASN	-	SEE REMARK 999	UNP B6K9M7
A	456	CYS	-	SEE REMARK 999	UNP B6K9M7
A	457	VAL	-	SEE REMARK 999	UNP B6K9M7
A	458	GLN	-	SEE REMARK 999	UNP B6K9M7
A	459	VAL	-	SEE REMARK 999	UNP B6K9M7
A	460	VAL	-	SEE REMARK 999	UNP B6K9M7
A	461	THR	-	SEE REMARK 999	UNP B6K9M7
A	462	GLU	-	SEE REMARK 999	UNP B6K9M7
A	463	CYS	-	SEE REMARK 999	UNP B6K9M7
A	464	VAL	-	SEE REMARK 999	UNP B6K9M7
A	465	GLY	-	SEE REMARK 999	UNP B6K9M7
A	466	ASN	-	SEE REMARK 999	UNP B6K9M7
A	467	GLN	-	SEE REMARK 999	UNP B6K9M7
A	468	PHE	-	SEE REMARK 999	UNP B6K9M7
A	469	ASP	-	SEE REMARK 999	UNP B6K9M7
A	470	GLN	-	SEE REMARK 999	UNP B6K9M7
A	471	THR	-	SEE REMARK 999	UNP B6K9M7
A	472	SER	-	SEE REMARK 999	UNP B6K9M7
A	473	LYS	-	SEE REMARK 999	UNP B6K9M7
A	474	ALA	-	SEE REMARK 999	UNP B6K9M7
A	475	CYS	-	SEE REMARK 999	UNP B6K9M7
A	476	CYS	-	SEE REMARK 999	UNP B6K9M7
A	477	THR	-	SEE REMARK 999	UNP B6K9M7
A	478	GLU	-	SEE REMARK 999	UNP B6K9M7
A	479	PRO	-	SEE REMARK 999	UNP B6K9M7
A	480	GLU	-	SEE REMARK 999	UNP B6K9M7
A	481	ALA	-	EXPRESSION TAG	UNP B6K9M7
A	482	ALA	-	EXPRESSION TAG	UNP B6K9M7
A	483	ALA	-	EXPRESSION TAG	UNP B6K9M7
A	484	LEU	-	EXPRESSION TAG	UNP B6K9M7
A	485	VAL	-	EXPRESSION TAG	UNP B6K9M7
A	486	PRO	-	EXPRESSION TAG	UNP B6K9M7
A	487	ARG	-	EXPRESSION TAG	UNP B6K9M7

- Molecule 2 is a protein called RHOPTRY NECK PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	30	Total	C	N	O	S	0	0	0
			229	143	36	45	5			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	995	GLY	-	EXPRESSION TAG	UNP B6KLP1
B	996	SER	-	EXPRESSION TAG	UNP B6KLP1
B	997	ALA	-	EXPRESSION TAG	UNP B6KLP1
B	998	SER	-	EXPRESSION TAG	UNP B6KLP1

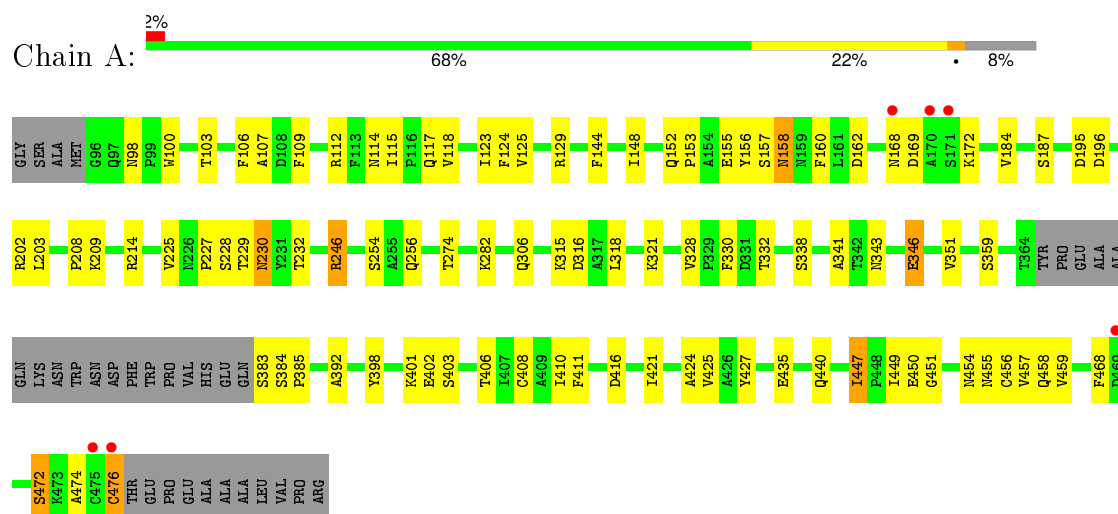
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	6	Total O 6 6	0	0
3	B	1	Total O 1 1	0	0

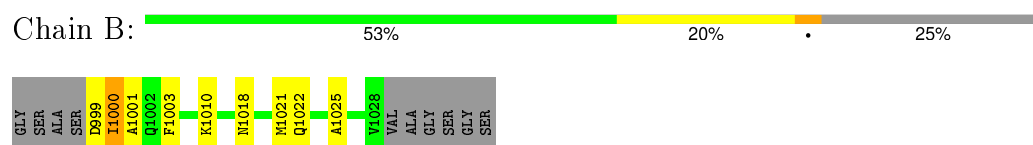
3 Residue-property plots [i](#)

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 ($\text{RSRZ} > 2$). 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 2: RHOPTRY NECK PROTEIN 2



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	49.37Å 124.18Å 171.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	85.96 – 3.10 85.96 – 3.10	Depositor EDS
% Data completeness (in resolution range)	96.8 (85.96-3.10) 96.8 (85.96-3.10)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.17 (at 3.13Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.205 , 0.263 0.206 , 0.265	Depositor DCC
R_{free} test set	994 reflections (11.42%)	DCC
Wilson B-factor (Å ²)	58.2	Xtriage
Anisotropy	0.052	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 49.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 9663 reflections (0.010%)	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	2981	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.58	0/2816	0.63	0/3836
2	B	0.52	0/232	0.61	0/313
All	All	0.57	0/3048	0.63	0/4149

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	2745	0	2589	72	0
2	B	229	0	218	9	0
3	A	6	0	0	2	0
3	B	1	0	0	0	0
All	All	2981	0	2807	77	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 13.

The worst 5 of 77 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:144:PHE:O	1:A:282:LYS:HE2	1.64	0.95

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:246:ARG:HH11	1:A:246:ARG:HG2	1.31	0.92
1:A:476:CYS:SG	1:A:476:CYS:O	2.42	0.78
1:A:343:ASN:OD1	1:A:346:GLU:HB2	1.89	0.73
2:B:1000:ILE:HG12	2:B:1001:ALA:N	2.06	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	359/396 (91%)	342 (95%)	16 (4%)	1 (0%)	46	80
2	B	28/40 (70%)	26 (93%)	2 (7%)	0	100	100
All	All	387/436 (89%)	368 (95%)	18 (5%)	1 (0%)	46	80

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	168	ASN

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	298/324 (92%)	281 (94%)	17 (6%)	25	62

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	26/31 (84%)	23 (88%)	3 (12%)	7	27
All	All	324/355 (91%)	304 (94%)	20 (6%)	23	59

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	351	VAL
1	A	384	SER
1	A	476	CYS
1	A	274	THR
1	A	346	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	455	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 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 ⓘ

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	363/396 (91%)	0.26	6 (1%) 73 52	25, 45, 81, 110	4 (1%)
2	B	30/40 (75%)	0.26	0 100 100	39, 57, 71, 86	0
All	All	393/436 (90%)	0.26	6 (1%) 76 58	25, 46, 81, 110	4 (1%)

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	476	CYS	9.1
1	A	168	ASN	6.4
1	A	475	CYS	3.7
1	A	469	ASP	2.4
1	A	171	SER	2.1

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.