Full wwPDB X-ray Structure Validation Report (i)

Sep 3, 2023 – 07:22 AM EDT

PDB ID	:	3SRI
Title	:	Crystal structure of Plasmodium falciparum AMA1 in complex with a 29aa PfRON2 peptide
Authors	:	Vulliez-Le Normand, B.; Saul, F.A.; Bentley, G.A.
Deposited on	:	2011-07-07
Resolution	:	1.60 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

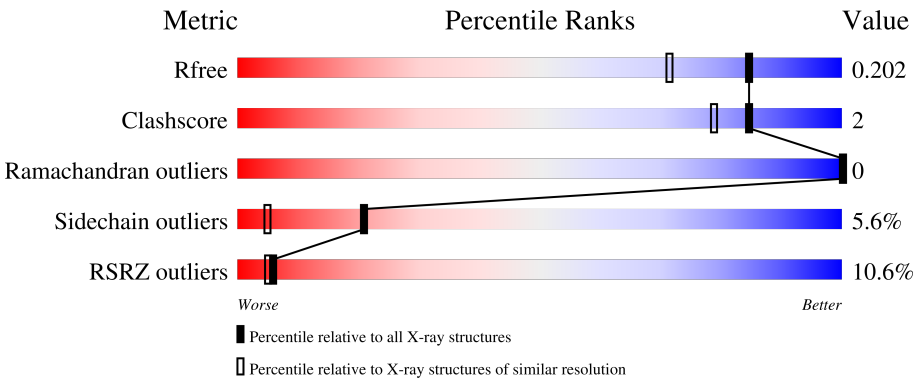
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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 >=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 <=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	381	6%; 68%; 7%; 25%
2	B	29	41%; 69%; 17%; 14%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2764 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	286	2309	1462	388	445	14	0	5	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	95	GLU	-	expression tag	UNP P50489
A	96	PHE	-	expression tag	UNP P50489
A	162	LYS	ASN	engineered mutation	UNP P50489
A	288	VAL	THR	engineered mutation	UNP P50489
A	373	ASP	SER	engineered mutation	UNP P50489
A	422	ASP	ASN	engineered mutation	UNP P50489
A	423	LYS	SER	engineered mutation	UNP P50489
A	443	MET	-	expression tag	UNP P50489
A	444	VAL	-	expression tag	UNP P50489
A	445	PRO	-	expression tag	UNP P50489
A	446	ARG	-	expression tag	UNP P50489
A	447	ALA	-	expression tag	UNP P50489
A	448	ALA	-	expression tag	UNP P50489
A	449	ALA	-	expression tag	UNP P50489
A	450	ALA	-	expression tag	UNP P50489
A	451	ALA	-	expression tag	UNP P50489
A	452	SER	-	expression tag	UNP P50489
A	453	PHE	-	expression tag	UNP P50489
A	454	LEU	-	expression tag	UNP P50489
A	455	GLU	-	expression tag	UNP P50489
A	456	GLN	-	expression tag	UNP P50489
A	457	LYS	-	expression tag	UNP P50489
A	458	LEU	-	expression tag	UNP P50489
A	459	ILE	-	expression tag	UNP P50489
A	460	SER	-	expression tag	UNP P50489
A	461	GLU	-	expression tag	UNP P50489
A	462	GLU	-	expression tag	UNP P50489

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	463	ASP	-	expression tag	UNP P50489
A	464	LEU	-	expression tag	UNP P50489
A	465	ASN	-	expression tag	UNP P50489
A	466	SER	-	expression tag	UNP P50489
A	467	ALA	-	expression tag	UNP P50489
A	468	VAL	-	expression tag	UNP P50489
A	469	ASP	-	expression tag	UNP P50489
A	470	HIS	-	expression tag	UNP P50489
A	471	HIS	-	expression tag	UNP P50489
A	472	HIS	-	expression tag	UNP P50489
A	473	HIS	-	expression tag	UNP P50489
A	474	HIS	-	expression tag	UNP P50489
A	475	HIS	-	expression tag	UNP P50489

- Molecule 2 is a protein called Rhoptry neck protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	25	190	118	34	35	3	0	1	0

- Molecule 3 is water.

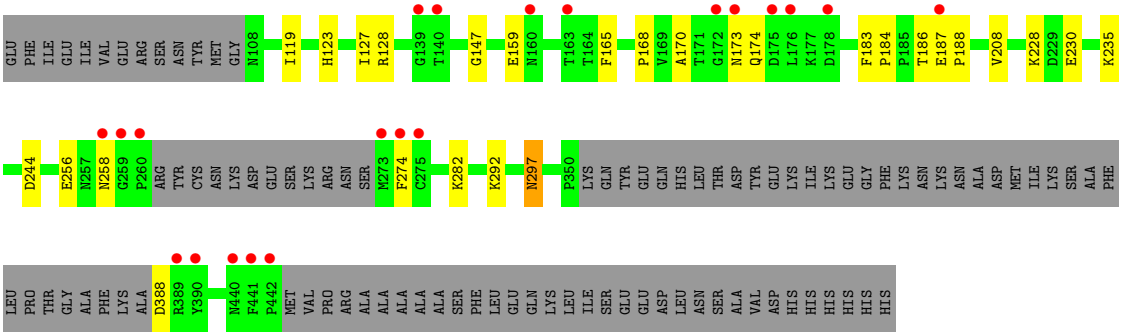
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	252	252	252	0	0
			Total	O		
3	B	13	13	13	0	0

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.

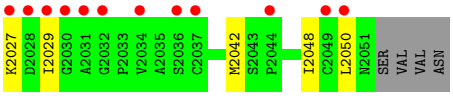
• Molecule 1: Apical membrane antigen 1

Chain A:



• Molecule 2: Rhoptry neck protein 2

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	70.72Å 38.14Å 72.08Å 90.00° 97.72° 90.00°	Depositor
Resolution (Å)	35.04 – 1.60 35.04 – 1.60	Depositor EDS
% Data completeness (in resolution range)	94.9 (35.04-1.60) 94.9 (35.04-1.60)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.88 (at 1.60Å)	Xtriage
Refinement program	BUSTER 2.9.3	Depositor
R, R_{free}	0.176 , 0.195 0.184 , 0.202	Depositor DCC
R_{free} test set	2462 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	19.0	Xtriage
Anisotropy	0.501	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 48.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.023 for l,-k,h	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2764	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.12% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0/2383	0.63	0/3229
2	B	0.41	0/193	0.61	0/260
All	All	0.50	0/2576	0.63	0/3489

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	2309	0	2185	11	0
2	B	190	0	191	4	0
3	A	252	0	0	0	0
3	B	13	0	0	0	0
All	All	2764	0	2376	11	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 (11) 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:297:ASN:HD22	1:A:297:ASN:H	1.43	0.65

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:168:PRO:HB3	1:A:184:PRO:HA	1.82	0.61
1:A:184:PRO:HG2	2:B:2050:LEU:HD22	1.83	0.60
1:A:119:ILE:HG21	1:A:127:ILE:HD11	1.94	0.49
1:A:170:ALA:HB1	1:A:174:GLN:HB3	1.95	0.48
1:A:188:PRO:HD2	2:B:2048:ILE:HG12	1.96	0.47
1:A:235:LYS:HB2	1:A:292:LYS:HD2	2.00	0.44
1:A:128:ARG:HH12	1:A:256:GLU:CD	2.22	0.43
1:A:123:HIS:CE1	1:A:147:GLY:HA3	2.53	0.43
1:A:208:VAL:HG21	2:B:2042:MET:SD	2.60	0.42
1:A:186:THR:HG22	2:B:2050:LEU:HD23	2.02	0.41

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	285/381 (75%)	271 (95%)	14 (5%)	0	100	100
2	B	23/29 (79%)	23 (100%)	0	0	100	100
All	All	308/410 (75%)	294 (96%)	14 (4%)	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	254/335 (76%)	239 (94%)	15 (6%)	19 4
2	B	22/25 (88%)	19 (86%)	3 (14%)	3 0
All	All	276/360 (77%)	258 (94%)	18 (6%)	21 3

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	159	GLU
1	A	165	PHE
1	A	173	ASN
1	A	183	PHE
1	A	187[A]	GLU
1	A	187[B]	GLU
1	A	228	LYS
1	A	230	GLU
1	A	244	ASP
1	A	258[A]	ASN
1	A	258[B]	ASN
1	A	274	PHE
1	A	282	LYS
1	A	297	ASN
1	A	388	ASP
2	B	2027[A]	LYS
2	B	2027[B]	LYS
2	B	2029	ILE

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

Mol	Chain	Res	Type
1	A	160	ASN
1	A	297	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 monosaccharides 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	286/381 (75%)	0.09	21 (7%)	15	13	13, 21, 60, 76	0
2	B	25/29 (86%)	1.81	12 (48%)	0	0	29, 47, 59, 76	0
All	All	311/410 (75%)	0.23	33 (10%)	6	5	13, 23, 60, 76	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	273	MET	8.4
1	A	259	GLY	8.2
1	A	260	PRO	8.1
1	A	390	TYR	4.5
2	B	2050	LEU	4.4
1	A	442	PRO	4.3
1	A	175	ASP	4.3
1	A	176	LEU	4.2
2	B	2034	VAL	3.9
1	A	172	GLY	3.8
2	B	2044	PRO	3.7
1	A	178	ASP	3.6
1	A	163	THR	3.5
2	B	2036	SER	3.4
1	A	258[A]	ASN	3.3
2	B	2028	ASP	3.2
1	A	440	ASN	3.1
1	A	274	PHE	3.0
1	A	160	ASN	3.0
2	B	2037	CYS	3.0
1	A	173	ASN	2.9
1	A	187[A]	GLU	2.9
2	B	2027[A]	LYS	2.8
1	A	441	PHE	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	389	ARG	2.6
2	B	2032	GLY	2.5
2	B	2030	GLY	2.3
1	A	275	CYS	2.2
2	B	2029	ILE	2.2
2	B	2049	CYS	2.1
2	B	2031	ALA	2.1
1	A	140	THR	2.0
1	A	139	GLY	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.