Full wwPDB X-ray Structure Validation Report (i)

Sep 28, 2024 – 10:12 PM EDT

PDB ID	:	1Z40
Title	:	AMA1 from Plasmodium falciparum
Authors	:	Bai, T.; Becker, M.; Gupta, A.; Strike, P.; Murphy, V.J.; Anders, R.F.; Batchelor, A.H.
Deposited on	:	2005-03-14
Resolution	:	1.90 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:

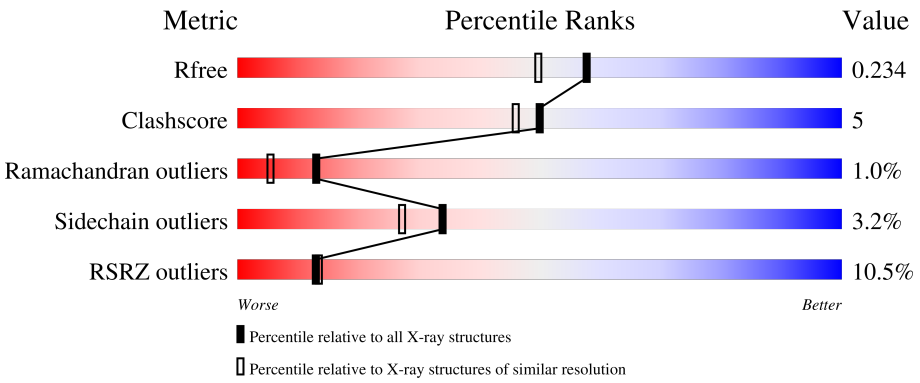
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.90 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	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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%

Mol	Chain	Length	Quality of chain
1	A	336	11% 81% 11% • 7%
1	E	336	9% 79% 11% • 9%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5136 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 precursor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	311	2461	1562	408	475	16	0	0	0
1	E	306	2421	1537	400	468	16	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	103	GLY	-	cloning artifact	UNP Q7KQK5
E	103	GLY	-	cloning artifact	UNP Q7KQK5

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
2	A	1	1	1	0	0

- Molecule 3 is water.

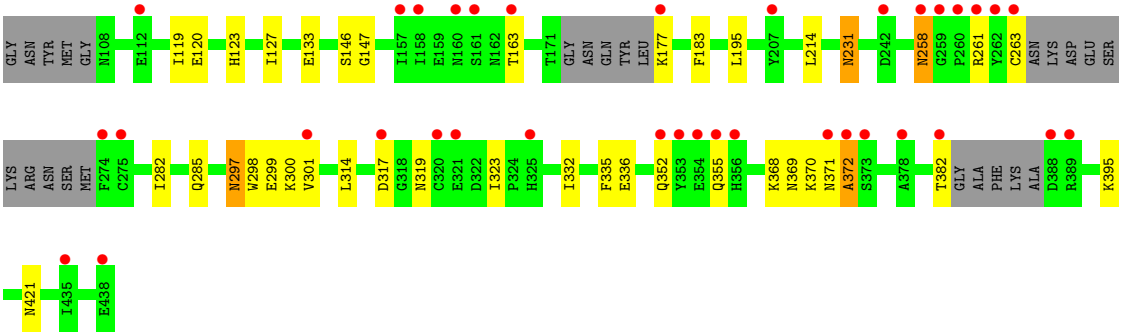
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	128	128	128	0	0
3	E	125	125	125	0	0

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. 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. 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.

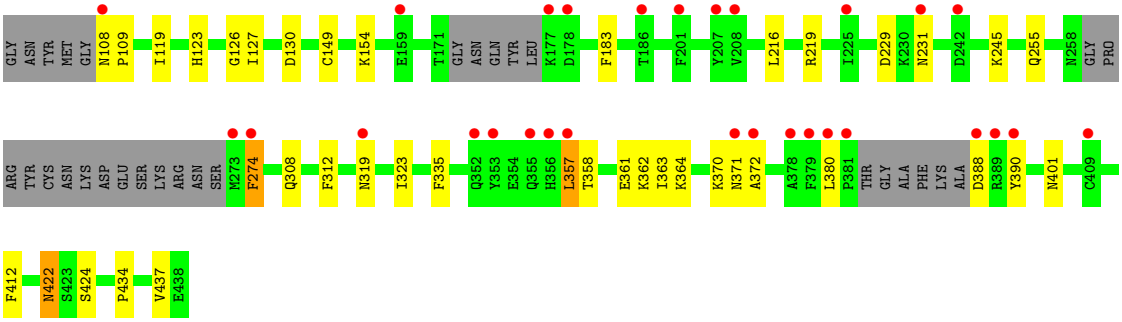
- Molecule 1: apical membrane antigen 1 precursor

Chain A:



- Molecule 1: apical membrane antigen 1 precursor

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	54.12Å 54.12Å 214.14Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.22 – 1.90 39.22 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.3 (39.22-1.90) 98.8 (39.22-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	10.79 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.192 , 0.236 0.190 , 0.234	Depositor DCC
R_{free} test set	2768 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	30.3	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 48.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.006 for -h,-k,l 0.043 for h,-h-k,-l 0.027 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5136	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.01% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: CL

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	0.69	0/2523	0.71	0/3421
1	E	0.71	0/2481	0.72	1/3361 (0.0%)
All	All	0.70	0/5004	0.71	1/6782 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	219	ARG	NE-CZ-NH2	-6.31	117.14	120.30

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	2461	0	2311	22	0
1	E	2421	0	2273	27	0
2	A	1	0	0	0	0
3	A	128	0	0	1	0
3	E	125	0	0	1	0
All	All	5136	0	4584	49	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 5.

All (49) 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:299:GLU:HG3	3:A:487:HOH:O	1.65	0.95
1:E:312:PHE:H	1:E:401:ASN:HD21	1.21	0.88
1:E:130:ASP:H	1:E:255:GLN:HE22	1.23	0.82
1:A:368:LYS:HZ1	1:A:369:ASN:HD21	1.34	0.76
1:E:123:HIS:HD2	1:E:149:CYS:H	1.36	0.74
1:A:368:LYS:NZ	1:A:369:ASN:HD21	1.89	0.69
1:E:229:ASP:OD1	1:E:231:ASN:ND2	2.26	0.68
1:A:297:ASN:HD22	1:A:297:ASN:H	1.44	0.65
1:E:422:ASN:HD22	1:E:424:SER:H	1.44	0.64
1:E:312:PHE:H	1:E:401:ASN:ND2	1.93	0.64
1:E:357:LEU:HD22	1:E:362:LYS:HG3	1.82	0.60
1:E:308:GLN:HG3	1:E:437:VAL:HG13	1.84	0.59
1:E:154:LYS:NZ	1:E:255:GLN:HE21	2.01	0.59
1:A:370:LYS:O	1:A:370:LYS:HG3	2.04	0.58
1:A:133:GLU:HG3	1:A:146:SER:CB	2.34	0.58
1:E:123:HIS:CD2	1:E:149:CYS:H	2.22	0.56
1:A:300:LYS:HG3	1:A:301:VAL:HG13	1.88	0.56
1:E:422:ASN:ND2	1:E:424:SER:H	2.06	0.53
1:A:119:ILE:HG21	1:A:127:ILE:HD11	1.89	0.53
1:E:363:ILE:HD13	1:E:380:LEU:HD11	1.90	0.53
1:A:258:ASN:HA	1:A:263:CYS:SG	2.50	0.52
1:E:422:ASN:HD22	1:E:422:ASN:C	2.14	0.51
1:E:312:PHE:N	1:E:401:ASN:HD21	2.02	0.50
1:A:368:LYS:HZ1	1:A:369:ASN:ND2	2.04	0.50
1:E:245:LYS:NZ	3:E:501:HOH:O	2.46	0.48
1:E:154:LYS:HZ3	1:E:255:GLN:HE21	1.59	0.48
1:E:335:PHE:CD1	1:E:434:PRO:HB2	2.49	0.47
1:E:335:PHE:HD1	1:E:434:PRO:HB2	1.79	0.47
1:E:308:GLN:HG3	1:E:437:VAL:CG1	2.44	0.47
1:A:332:ILE:HG22	1:A:336:GLU:HG3	1.96	0.46
1:E:154:LYS:HZ3	1:E:255:GLN:NE2	2.13	0.46
1:A:258:ASN:HD22	1:A:258:ASN:N	2.14	0.46
1:E:370:LYS:O	1:E:372:ALA:N	2.49	0.46
1:E:364:LYS:HD2	1:E:390:TYR:HE1	1.81	0.45
1:A:123:HIS:CE1	1:A:147:GLY:HA3	2.52	0.45
1:E:123:HIS:HE1	1:E:126:GLY:O	1.99	0.45
1:E:323:ILE:HD13	1:E:412:PHE:CZ	2.52	0.43
1:A:120:GLU:OE1	1:A:395:LYS:HE3	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:214:LEU:HD12	1:A:298:TRP:CZ2	2.54	0.42
1:A:314:LEU:HG	1:A:323:ILE:HD13	2.02	0.42
1:E:358:THR:OG1	1:E:361:GLU:HG2	2.20	0.42
1:A:297:ASN:OD1	1:A:300:LYS:HE2	2.19	0.42
1:A:231:ASN:H	1:A:231:ASN:HD22	1.67	0.41
1:A:133:GLU:HG3	1:A:146:SER:HB3	2.01	0.41
1:A:163:THR:HG21	1:A:177:LYS:O	2.21	0.41
1:A:282:ILE:HG13	1:A:285:GLN:NE2	2.36	0.41
1:A:371:ASN:O	1:A:372:ALA:HB3	2.20	0.41
1:E:119:ILE:HG21	1:E:127:ILE:HD11	2.02	0.41
1:E:108:ASN:HA	1:E:109:PRO:HD3	1.95	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	303/336 (90%)	287 (95%)	12 (4%)	4 (1%)	10 3
1	E	298/336 (89%)	287 (96%)	9 (3%)	2 (1%)	19 11
All	All	601/672 (89%)	574 (96%)	21 (4%)	6 (1%)	13 5

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	261	ARG
1	A	355	GLN
1	E	274	PHE
1	A	352	GLN
1	E	371	ASN
1	A	372	ALA

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	267/296 (90%)	257 (96%)	10 (4%)	29	22
1	E	261/296 (88%)	254 (97%)	7 (3%)	40	34
All	All	528/592 (89%)	511 (97%)	17 (3%)	34	27

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	183	PHE
1	A	195	LEU
1	A	231	ASN
1	A	258	ASN
1	A	297	ASN
1	A	317	ASP
1	A	319	ASN
1	A	335	PHE
1	A	382	THR
1	A	421	ASN
1	E	183	PHE
1	E	216	LEU
1	E	274	PHE
1	E	319	ASN
1	E	357	LEU
1	E	388	ASP
1	E	422	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	231	ASN
1	A	233	ASN
1	A	257	ASN
1	A	258	ASN
1	A	285	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	369	ASN
1	A	421	ASN
1	E	123	HIS
1	E	223	ASN
1	E	231	ASN
1	E	255	GLN
1	E	257	ASN
1	E	319	ASN
1	E	401	ASN
1	E	422	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)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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)

Warning: The R factor obtained from EDS is 0.2501,which does not match the depositor's R factor of 0.19241. Please interpret the results in this section carefully.

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	311/336 (92%)	0.98	36 (11%)	11	11	29, 37, 51, 61	0
1	E	306/336 (91%)	0.88	29 (9%)	15	16	30, 37, 47, 55	0
All	All	617/672 (91%)	0.93	65 (10%)	13	13	29, 37, 48, 61	0

All (65) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	353	TYR	5.3
1	A	372	ALA	5.3
1	A	262	TYR	5.1
1	A	352	GLN	4.9
1	A	263	CYS	4.8
1	E	381	PRO	4.7
1	A	353	TYR	4.3
1	A	274	PHE	3.8
1	A	438	GLU	3.8
1	E	372	ALA	3.7
1	E	231	ASN	3.6
1	E	357	LEU	3.5
1	E	352	GLN	3.5
1	A	163	THR	3.5
1	A	388	ASP	3.4
1	A	260	PRO	3.4
1	A	325	HIS	3.3
1	E	273	MET	3.3
1	A	161	SER	3.2
1	E	319	ASN	3.1
1	A	259	GLY	3.1
1	A	261	ARG	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	207	TYR	2.7
1	A	354	GLU	2.7
1	A	355	GLN	2.7
1	A	275	CYS	2.6
1	A	160	ASN	2.6
1	E	186	THR	2.6
1	E	380	LEU	2.6
1	A	242	ASP	2.6
1	E	390	TYR	2.5
1	E	379	PHE	2.5
1	E	371	ASN	2.5
1	A	389	ARG	2.4
1	E	378	ALA	2.4
1	A	321	GLU	2.4
1	E	208	VAL	2.3
1	A	382	THR	2.3
1	E	108	ASN	2.3
1	E	201	PHE	2.3
1	E	388	ASP	2.3
1	E	178	ASP	2.3
1	E	389	ARG	2.2
1	E	242	ASP	2.2
1	E	355	GLN	2.2
1	E	177	LYS	2.2
1	A	207	TYR	2.2
1	A	158	ILE	2.2
1	A	373	SER	2.2
1	A	371	ASN	2.2
1	E	159	GLU	2.2
1	A	356	HIS	2.2
1	A	317	ASP	2.1
1	A	258	ASN	2.1
1	A	301	VAL	2.1
1	E	409	CYS	2.1
1	A	378	ALA	2.1
1	A	435	ILE	2.1
1	A	320	CYS	2.1
1	E	356	HIS	2.1
1	E	225	ILE	2.1
1	A	177	LYS	2.1
1	A	112	GLU	2.1
1	A	157	ILE	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	274	PHE	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)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CL	A	1	1/1	0.93	0.21	53,53,53,53	0

6.5 Other polymers (i)

There are no such residues in this entry.