



Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 09:06 pm BST

PDB ID : 4IOD
Title : Preliminary structural investigations of a malarial protein secretion system.
Authors : Egea, P.F.
Deposited on : 2013-01-07
Resolution : 1.80 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

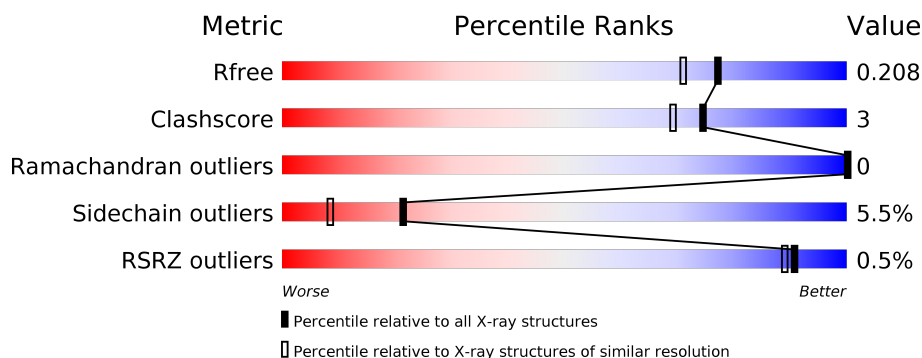
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 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	156	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 87%, grey 11%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 87% 7% • 6% </div> </div>
1	B	156	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 83%, yellow 9%, orange 6%, grey 6%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> 83% 9% • 6% </div> </div>
1	C	156	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 10%, green 82%, grey 8%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 82% 10% • 6% </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	C	201	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3895 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 Malarial CLP B2 atpase/HSP101 protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	B	147	Total	C	N	O	5	7	0
			1171	748	198	225			
1	A	147	Total	C	N	O	0	5	0
			1165	740	199	226			
1	C	146	Total	C	N	O	0	7	0
			1173	744	204	225			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	EXPRESSION TAG	UNP Q8IIJ8
B	153	LEU	-	EXPRESSION TAG	UNP Q8IIJ8
B	154	VAL	-	EXPRESSION TAG	UNP Q8IIJ8
B	155	PRO	-	EXPRESSION TAG	UNP Q8IIJ8
B	156	ARG	-	EXPRESSION TAG	UNP Q8IIJ8
A	1	MET	-	EXPRESSION TAG	UNP Q8IIJ8
A	153	LEU	-	EXPRESSION TAG	UNP Q8IIJ8
A	154	VAL	-	EXPRESSION TAG	UNP Q8IIJ8
A	155	PRO	-	EXPRESSION TAG	UNP Q8IIJ8
A	156	ARG	-	EXPRESSION TAG	UNP Q8IIJ8
C	1	MET	-	EXPRESSION TAG	UNP Q8IIJ8
C	153	LEU	-	EXPRESSION TAG	UNP Q8IIJ8
C	154	VAL	-	EXPRESSION TAG	UNP Q8IIJ8
C	155	PRO	-	EXPRESSION TAG	UNP Q8IIJ8
C	156	ARG	-	EXPRESSION TAG	UNP Q8IIJ8

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).

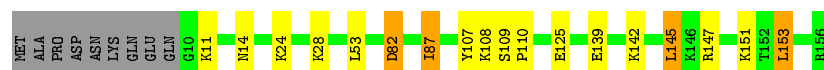


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		

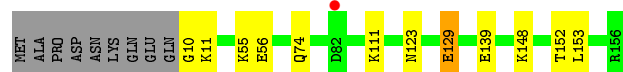
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	147	Total	O	0	0
			147	147		
3	A	122	Total	O	0	0
			122	122		
3	C	102	Total	O	0	0
			102	102		

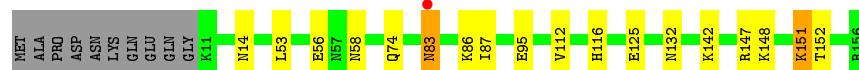
- Molecule 1: Malarial CLP B2 atpase/HSP101 protein



- Molecule 1: Malarial CLP B2 atpase/HSP101 protein



- Molecule 1: Malarial CLP B2 atpase/HSP101 protein



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	89.18Å 141.01Å 91.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.59 – 1.80 75.37 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.7 (44.59-1.80) 99.7 (75.37-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 1.80Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.170 , 0.212 0.167 , 0.208	Depositor DCC
R_{free} test set	5334 reflections (9.97%)	wwPDB-VP
Wilson B-factor (Å ²)	21.5	Xtriage
Anisotropy	0.116	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 46.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3895	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.66% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.54	3/1189 (0.3%)	0.53	0/1607
1	B	0.50	1/1204 (0.1%)	0.62	2/1628 (0.1%)
1	C	0.58	1/1206 (0.1%)	0.59	0/1629
All	All	0.54	5/3599 (0.1%)	0.58	2/4864 (0.0%)

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	A	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	55	LYS	C-N	7.21	1.50	1.34
1	A	129	GLU	CD-OE1	-6.45	1.18	1.25
1	C	56	GLU	CD-OE1	-6.34	1.18	1.25
1	B	110	PRO	N-CD	5.46	1.55	1.47
1	A	56	GLU	C-N	-5.38	1.21	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	109[A]	SER	C-N-CD	5.21	139.35	128.40
1	B	109[B]	SER	C-N-CD	5.21	139.35	128.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	139[B]	GLU	Mainchain

5.2 Too-close contacts ⓘ

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	1165	0	1197	6	0
1	B	1171	0	1218	10	0
1	C	1173	0	1218	8	1
2	A	5	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	2	0
3	A	122	0	0	4	0
3	B	147	0	0	1	0
3	C	102	0	0	2	0
All	All	3895	0	3633	23	1

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 3.

All (23) 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:C:151:LYS:HE3	2:C:201:SO4:O1	1.77	0.85
1:B:147:ARG:NH1	3:B:417:HOH:O	2.25	0.69
1:A:10:GLY:O	1:A:111:LYS:NZ	2.23	0.69
1:C:151:LYS:CE	2:C:201:SO4:O1	2.45	0.64
1:C:132[A]:ASN:OD1	3:C:377:HOH:O	2.18	0.55
1:C:148:LYS:HE3	1:C:152:THR:HG21	1.90	0.53
1:B:107:TYR:O	1:B:108:LYS:HB2	2.08	0.53
1:B:147:ARG:NH1	1:B:151:LYS:HE3	2.24	0.52
1:B:107:TYR:CE2	1:B:145[B]:LEU:HD13	2.46	0.51
1:C:83:ASN:OD1	1:C:83:ASN:N	2.44	0.49
1:A:123:ASN:HB2	3:A:420:HOH:O	2.14	0.48
1:C:147[B]:ARG:NH1	3:C:367:HOH:O	2.47	0.47
1:A:123:ASN:ND2	3:A:399:HOH:O	2.37	0.47
1:B:153:LEU:HD12	1:C:74:GLN:HB3	1.97	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:152[B]:THR:HG21	3:A:371:HOH:O	2.14	0.46
1:B:82:ASP:N	1:B:82:ASP:OD1	2.39	0.46
1:A:148:LYS:O	1:A:152[B]:THR:HG22	2.17	0.45
1:B:28:LYS:HG2	1:B:87[B]:ILE:HD11	1.98	0.45
1:C:112:VAL:HG13	1:C:116:HIS:HB2	1.99	0.44
1:A:129:GLU:HG2	3:A:422:HOH:O	2.18	0.43
1:B:24:LYS:HG2	1:B:87[A]:ILE:HG12	1.99	0.43
1:B:142:LYS:HB3	1:B:142:LYS:HE3	1.80	0.43
1:B:11:LYS:HB2	1:B:11:LYS:HE3	1.62	0.42

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:58:ASN:O	1:C:147[A]:ARG:NH1[3_554]	2.13	0.07

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	150/156 (96%)	150 (100%)	0	0	100	100
1	B	152/156 (97%)	151 (99%)	1 (1%)	0	100	100
1	C	151/156 (97%)	151 (100%)	0	0	100	100
All	All	453/468 (97%)	452 (100%)	1 (0%)	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	126/132 (96%)	123 (98%)	3 (2%)	49	36
1	B	128/132 (97%)	116 (91%)	12 (9%)	8	2
1	C	129/132 (98%)	120 (93%)	9 (7%)	15	5
All	All	383/396 (97%)	359 (94%)	24 (6%)	21	6

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	14	ASN
1	B	53[A]	LEU
1	B	53[B]	LEU
1	B	82	ASP
1	B	87[A]	ILE
1	B	87[B]	ILE
1	B	125	GLU
1	B	139[A]	GLU
1	B	139[B]	GLU
1	B	145[A]	LEU
1	B	145[B]	LEU
1	B	153	LEU
1	A	11	LYS
1	A	74	GLN
1	A	153	LEU
1	C	14	ASN
1	C	53	LEU
1	C	83	ASN
1	C	86	LYS
1	C	87	ILE
1	C	95	GLU
1	C	125	GLU
1	C	142	LYS
1	C	151	LYS

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

Mol	Chain	Res	Type
1	B	128	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](#)

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	SO4	B	201	-	4,4,4	0.17	0	6,6,6	0.33	0
2	SO4	A	201	-	4,4,4	0.13	0	6,6,6	0.13	0
2	SO4	C	201	-	4,4,4	0.12	0	6,6,6	0.14	0

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.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	201	SO4	2	0

5.7 Other polymers

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	147/156 (94%)	-0.24	1 (0%) 87 86	13, 21, 37, 55	0
1	B	147/156 (94%)	-0.24	0 100 100	12, 20, 37, 48	0
1	C	146/156 (93%)	-0.20	1 (0%) 87 86	13, 23, 42, 67	0
All	All	440/468 (94%)	-0.23	2 (0%) 91 89	12, 21, 41, 67	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	82	ASP	3.3
1	C	83	ASN	3.0

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 carbohydrates 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	SO4	B	201	5/5	0.86	0.20	46,47,52,52	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	SO4	C	201	5/5	0.92	0.21	76,76,76,77	0
2	SO4	A	201	5/5	0.98	0.09	33,33,34,34	0

6.5 Other polymers [i](#)

There are no such residues in this entry.