Full wwPDB X-ray Structure Validation Report ⓘ

May 5, 2025 – 10:57 AM EDT

PDB ID : 9BPS / pdb_00009bps
Title : Plasmodium falciparum apicoplast DNA polymerase mutant - K417M
Authors : Ung, A.R.; Honzatko, R.B.; Nelson, S.W.
Deposited on : 2024-05-08
Resolution : 2.80 Å(reported)

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

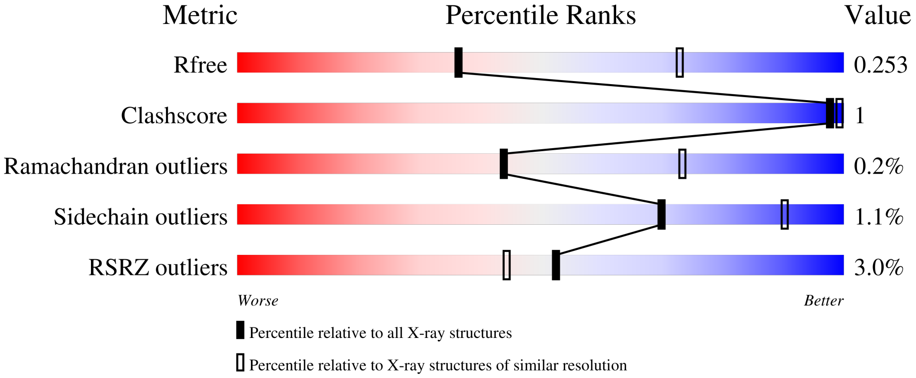
MolProbity : 4-5-2 with Phenix2.0rc1
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.006 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.43.1

1 Overall quality at a glance ⓘ

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

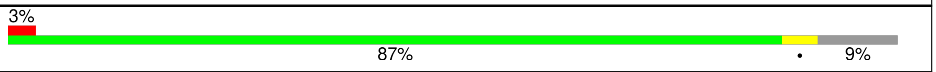
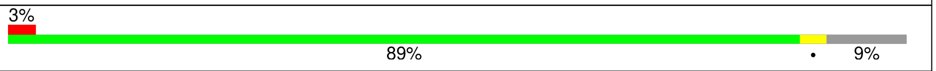
The reported resolution of this entry is 2.80 Å.

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}	164625	3657 (2.80-2.80)
Clashscore	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)
RSRZ outliers	164620	3659 (2.80-2.80)

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 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	628	
1	B	628	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 19468 atoms, of which 9758 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 Plastid replication-repair enzyme.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	572	9629	3077	4858	795	882	17	0	3	0
1	B	574	9650	3083	4870	796	884	17	0	2	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	82	ASN	ASP	engineered mutation	UNP Q8ILY1
A	84	GLN	GLU	engineered mutation	UNP Q8ILY1
A	417	MET	LYS	engineered mutation	UNP Q8ILY1
B	82	ASN	ASP	engineered mutation	UNP Q8ILY1
B	84	GLN	GLU	engineered mutation	UNP Q8ILY1
B	417	MET	LYS	engineered mutation	UNP Q8ILY1

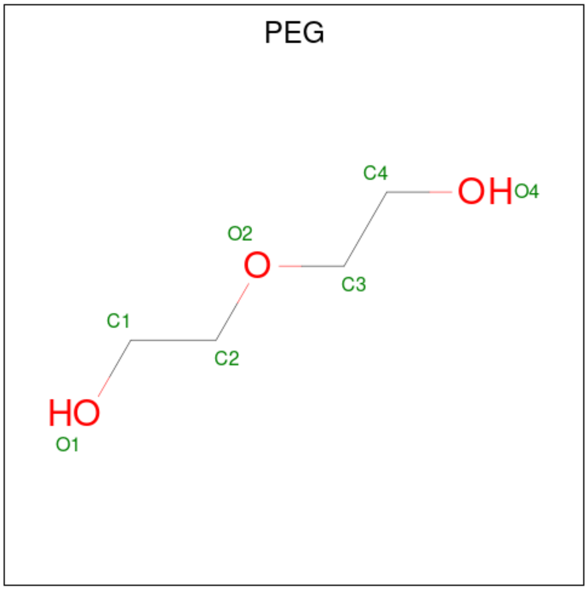
- Molecule 2 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Na		
2	A	1	1	1	0	0
2	B	1	1	1	0	0

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
3	A	2	2	2	0	0
3	B	2	2	2	0	0

- Molecule 4 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
4	A	1	17	4	10	3	0	0
4	B	1	17	4	10	3	0	0
4	B	1	17	4	10	3	0	0

- Molecule 5 is water.

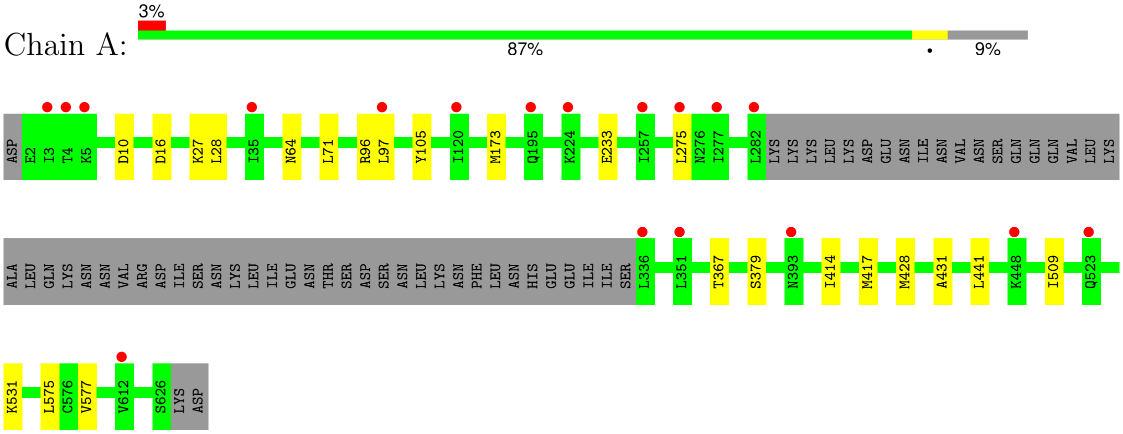
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	63	63	63	0	0
5	B	69	69	69	0	0

3 Residue-property plots ⓘ

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.

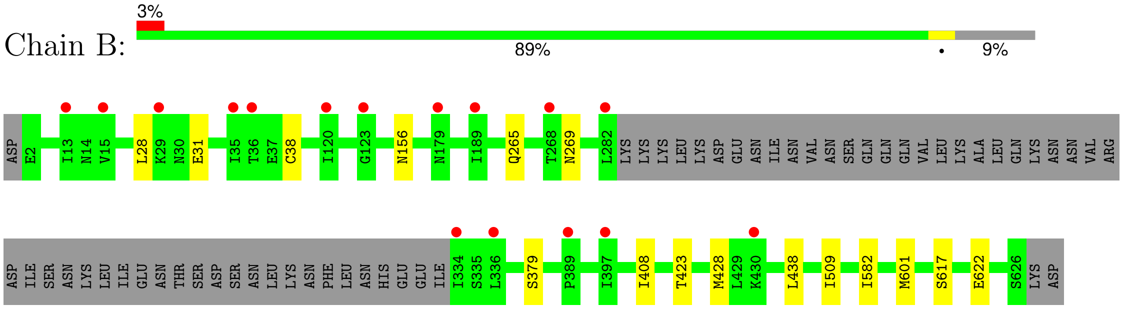
- Molecule 1: Plastid replication-repair enzyme

Chain A:



- Molecule 1: Plastid replication-repair enzyme

Chain B:



4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	144.93Å 144.93Å 166.38Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.44 – 2.80 47.44 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.2 (47.44-2.80) 99.2 (47.44-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.29 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.231 , 0.253 0.228 , 0.253	Depositor DCC
R_{free} test set	46632 reflections (4.11%)	wwPDB-VP
Wilson B-factor (Å^2)	78.6	Xtriage
Anisotropy	0.162	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 51.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.034 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	19468	wwPDB-VP
Average B, all atoms (Å^2)	96.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.80% 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 ⓘ

5.1 Standard geometry ⓘ

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

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	0.22	0/4863	0.41	0/6548
1	B	0.21	0/4869	0.39	0/6556
All	All	0.22	0/9732	0.40	0/13104

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 ⓘ

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	4771	4858	4862	9	0
1	B	4780	4870	4872	4	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
4	A	7	10	10	0	0
4	B	14	20	20	0	0
5	A	63	0	0	0	0
5	B	69	0	0	0	0
All	All	9710	9758	9764	13	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 1.

All (13) 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:10:ASP:OD1	1:A:64:ASN:ND2	2.43	0.51
1:A:27:LYS:O	1:A:105:TYR:OH	2.28	0.49
1:B:428:MET:HE1	1:B:509:ILE:HD11	1.96	0.48
1:A:96:ARG:HG3	1:A:97:LEU:HD12	1.97	0.46
1:A:431:ALA:CB	1:A:441:LEU:HD22	2.46	0.46
1:B:265:GLN:O	1:B:269:ASN:ND2	2.49	0.46
1:A:575:LEU:HD21	1:A:577:VAL:HG22	1.98	0.45
1:B:423:THR:HG21	1:B:509:ILE:HD13	1.99	0.44
1:A:28:LEU:HD21	1:A:71:LEU:CB	2.50	0.42
1:A:428:MET:HE1	1:A:509:ILE:HD11	2.00	0.42
1:B:408:ILE:HG22	1:B:582:ILE:HG12	2.01	0.42
1:A:233:GLU:O	1:A:531:LYS:NZ	2.45	0.41
1:A:414:ILE:HA	1:A:417:MET:HE3	2.03	0.41

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	571/628 (91%)	553 (97%)	17 (3%)	1 (0%)	44	73
1	B	572/628 (91%)	556 (97%)	15 (3%)	1 (0%)	44	73
All	All	1143/1256 (91%)	1109 (97%)	32 (3%)	2 (0%)	44	73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	16	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	156	ASN

5.3.2 Protein sidechains ⓘ

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	539/591 (91%)	535 (99%)	4 (1%)	81	94
1	B	540/591 (91%)	532 (98%)	8 (2%)	60	86
All	All	1079/1182 (91%)	1067 (99%)	12 (1%)	70	90

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	173	MET
1	A	275	LEU
1	A	367	THR
1	A	379	SER
1	B	28	LEU
1	B	31	GLU
1	B	38	CYS
1	B	379	SER
1	B	438	LEU
1	B	601	MET
1	B	617	SER
1	B	622	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	114	ASN
1	A	116	ASN
1	A	139	ASN
1	A	172	ASN
1	B	114	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	187	ASN
1	B	194	GLN
1	B	196	ASN
1	B	338	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 6 are monoatomic - leaving 3 for Mogul analysis.

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	$\#\|Z\| > 2$
4	PEG	B	705	-	6,6,6	0.16	0	5,5,5	0.14	0
4	PEG	A	704	-	6,6,6	0.24	0	5,5,5	0.15	0
4	PEG	B	701	-	6,6,6	0.21	0	5,5,5	0.11	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	PEG	B	705	-	-	1/4/4/4	-
4	PEG	A	704	-	-	0/4/4/4	-
4	PEG	B	701	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	705	PEG	O1-C1-C2-O2
4	B	701	PEG	C4-C3-O2-C2
4	B	701	PEG	C1-C2-O2-C3
4	B	701	PEG	O2-C3-C4-O4

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	572/628 (91%)	0.15	18 (3%)	51	43	39, 91, 159, 180	3 (0%)
1	B	574/628 (91%)	0.13	16 (2%)	55	46	48, 91, 156, 178	2 (0%)
All	All	1146/1256 (91%)	0.14	34 (2%)	52	44	39, 91, 158, 180	5 (0%)

All (34) RSRZ outliers are listed below:

Mol	**Chain**	**Res**	**Type**	**RSRZ**
1	B	35	ILE	5.9
1	B	179	ASN	3.5
1	A	277	ILE	3.4
1	A	282	LEU	3.3
1	A	35	ILE	3.2
1	A	224	LYS	2.9
1	B	268	THR	2.7
1	B	282	LEU	2.7
1	B	15	VAL	2.6
1	B	336	LEU	2.6
1	B	430[A]	LYS	2.6
1	A	275	LEU	2.6
1	A	448[A]	LYS	2.6
1	A	393	ASN	2.5
1	A	612	VAL	2.5
1	A	3	ILE	2.5
1	B	189	ILE	2.5
1	A	120	ILE	2.4
1	B	334	ILE	2.4
1	B	120	ILE	2.3
1	B	123	GLY	2.2
1	B	29	LYS	2.2
1	A	257	ILE	2.2
1	A	336	LEU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	389	PRO	2.2
1	B	13	ILE	2.2
1	A	97	LEU	2.2
1	B	397	ILE	2.2
1	B	36	THR	2.1
1	A	195	GLN	2.1
1	A	5	LYS	2.1
1	A	4	THR	2.0
1	A	351	LEU	2.0
1	A	523[A]	GLN	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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
4	PEG	B	701	7/7	0.78	0.20	84,108,131,131	0
2	NA	B	702	1/1	0.81	0.11	68,68,68,68	0
3	CL	B	703	1/1	0.84	0.14	71,71,71,71	0
4	PEG	A	704	7/7	0.85	0.19	72,93,125,125	0
4	PEG	B	705	7/7	0.90	0.17	82,102,124,124	0
3	CL	B	704	1/1	0.94	0.12	65,65,65,65	0
3	CL	A	702	1/1	0.94	0.12	68,68,68,68	0
3	CL	A	703	1/1	0.96	0.26	66,66,66,66	0
2	NA	A	701	1/1	0.96	0.07	61,61,61,61	0

6.5 Other polymers ⓘ

There are no such residues in this entry.