Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 05:50 am BST

PDB ID	:	6CA8
Title	:	Crystal structure of Plasmodium falciparum topoisomerase II DNA-binding, cleavage and re-ligation domain
Authors	:	Kumar, S.; Kandavelu, P.; Rathod, P.K.
Deposited on	:	2018-01-29
Resolution	:	2.33 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:

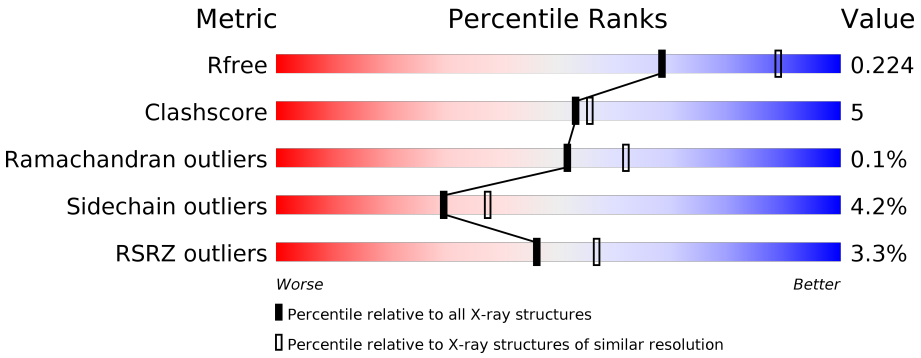
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.33 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	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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 <=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	751	3% 81% 15%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6462 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 DNA topoisomerase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	731	Total	C	N	O	S	0	0	0
			5926	3800	977	1127	22			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	468	MET	-	initiating methionine	UNP Q7Z2D0
A	1213	HIS	-	expression tag	UNP Q7Z2D0
A	1214	HIS	-	expression tag	UNP Q7Z2D0
A	1215	HIS	-	expression tag	UNP Q7Z2D0
A	1216	HIS	-	expression tag	UNP Q7Z2D0
A	1217	HIS	-	expression tag	UNP Q7Z2D0
A	1218	HIS	-	expression tag	UNP Q7Z2D0

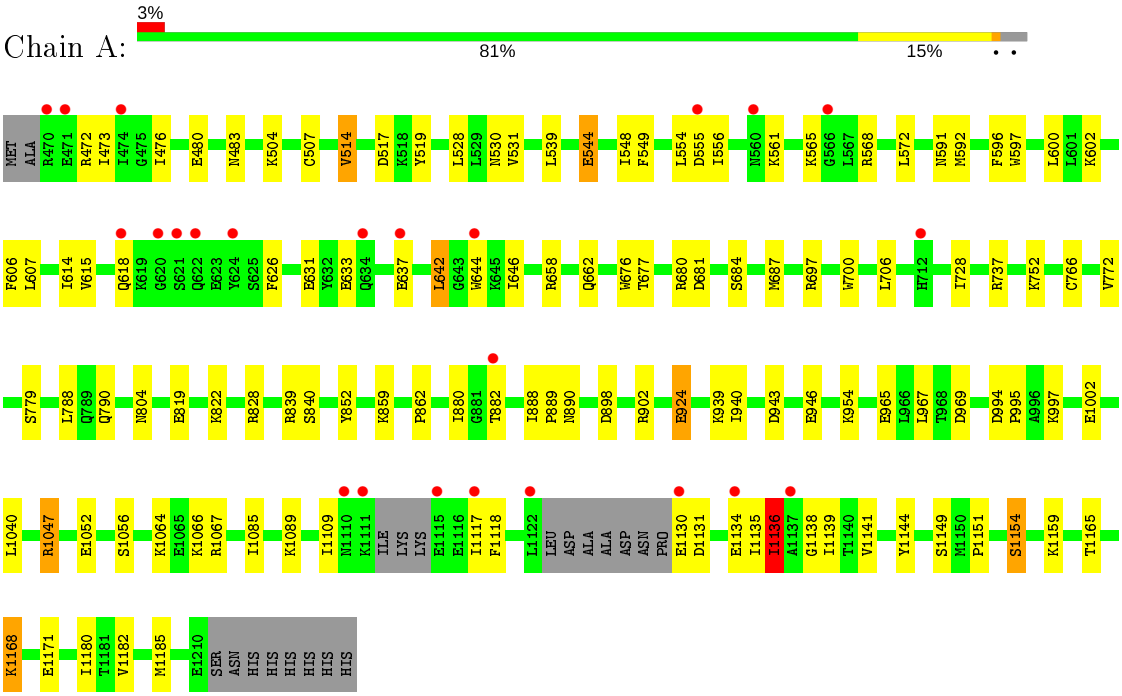
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	536	Total	O	0	0
			536	536		

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 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: DNA topoisomerase 2



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	83.03Å 104.48Å 114.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.22 – 2.33 44.22 – 2.33	Depositor EDS
% Data completeness (in resolution range)	98.0 (44.22-2.33) 98.0 (44.22-2.33)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.58 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.188 , 0.223 0.190 , 0.224	Depositor DCC
R_{free} test set	2135 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	21.6	Xtriage
Anisotropy	0.463	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 38.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6462	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% 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

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	0.32	0/6042	0.58	1/8141 (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	2

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	A	706	LEU	CA-CB-CG	6.96	131.30	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1136	ILE	Peptide
1	A	642	LEU	Peptide

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	5926	0	5853	64	0
2	A	536	0	0	5	0
All	All	6462	0	5853	64	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 (64) 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:1066:LYS:NZ	1:A:1171:GLU:OE1	1.94	0.99
1:A:1165:THR:HA	1:A:1168:LYS:NZ	1.92	0.84
1:A:642:LEU:H	1:A:644:TRP:HD1	1.33	0.76
1:A:924:GLU:OE2	2:A:1301:HOH:O	2.03	0.76
1:A:939:LYS:HD2	1:A:1002:GLU:OE2	1.89	0.72
1:A:1165:THR:HA	1:A:1168:LYS:HZ3	1.56	0.69
1:A:633:GLU:O	1:A:637:GLU:HG2	1.94	0.68
1:A:561:LYS:NZ	1:A:565:LYS:O	2.17	0.67
1:A:994:ASP:HB3	1:A:997:LYS:HB2	1.76	0.66
1:A:539:LEU:HD21	1:A:596:PHE:CZ	2.31	0.65
1:A:1165:THR:HA	1:A:1168:LYS:HZ2	1.60	0.65
1:A:677:THR:N	1:A:681:ASP:OD2	2.25	0.64
1:A:790:GLN:NE2	2:A:1306:HOH:O	2.30	0.64
1:A:680:ARG:NH2	2:A:1307:HOH:O	2.31	0.62
1:A:539:LEU:HD21	1:A:596:PHE:HZ	1.69	0.57
1:A:1149:SER:OG	2:A:1302:HOH:O	2.17	0.57
1:A:1067:ARG:HD3	1:A:1109:ILE:HG23	1.87	0.56
1:A:600:LEU:HD13	1:A:606:PHE:CZ	2.40	0.56
1:A:1151:PRO:O	1:A:1154:SER:OG	2.20	0.56
1:A:554:LEU:HD22	1:A:597:TRP:CE2	2.41	0.56
1:A:1085:ILE:O	1:A:1089:LYS:NZ	2.32	0.55
1:A:852:TYR:HA	1:A:862:PRO:HA	1.89	0.55
1:A:1052:GLU:HG3	1:A:1182:VAL:HG11	1.88	0.55
1:A:1135:ILE:HG22	1:A:1136:ILE:HG22	1.90	0.54
1:A:1134:GLU:OE1	1:A:1138:GLY:HA2	2.07	0.54
1:A:483:ASN:HB2	1:A:517:ASP:HA	1.88	0.54
1:A:556:ILE:H	1:A:556:ILE:HD12	1.73	0.54
1:A:473:ILE:HG21	1:A:476:ILE:HD12	1.90	0.53
1:A:626:PHE:HD1	1:A:631:GLU:HG2	1.72	0.53
1:A:602:LYS:HE3	1:A:676:TRP:NE1	2.23	0.53
1:A:700:TRP:CE2	1:A:728:ILE:HD12	2.43	0.53
1:A:804:ASN:ND2	1:A:890:ASN:HB3	2.25	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:819:GLU:O	1:A:822:LYS:HE3	2.11	0.50
1:A:1130:GLU:O	1:A:1131:ASP:HB2	2.10	0.50
1:A:943:ASP:OD1	1:A:995:PRO:HG3	2.12	0.50
1:A:549:PHE:CE1	1:A:597:TRP:HH2	2.31	0.49
1:A:1136:ILE:O	1:A:1139:ILE:HB	2.13	0.48
1:A:1165:THR:O	1:A:1168:LYS:HG3	2.13	0.48
1:A:614:ILE:HG23	1:A:615:VAL:HG23	1.95	0.48
1:A:480:GLU:OE2	1:A:504:LYS:HE3	2.15	0.47
1:A:1180:ILE:HD13	1:A:1185:MET:HG3	1.97	0.46
1:A:554:LEU:HD22	1:A:597:TRP:CD2	2.51	0.46
1:A:898:ASP:HB3	1:A:902:ARG:NH1	2.31	0.46
1:A:544:GLU:O	1:A:548:ILE:HG13	2.16	0.45
1:A:473:ILE:HD11	1:A:568:ARG:CZ	2.47	0.45
1:A:859:LYS:HB2	1:A:1118:PHE:CG	2.53	0.44
1:A:1154:SER:HA	1:A:1159:LYS:HG2	1.98	0.44
1:A:822:LYS:NZ	2:A:1329:HOH:O	2.50	0.44
1:A:642:LEU:N	1:A:644:TRP:HD1	2.10	0.43
1:A:531:VAL:HG21	1:A:592:MET:HA	2.00	0.43
1:A:658:ARG:O	1:A:662:GLN:HG3	2.18	0.43
1:A:940:ILE:HD11	1:A:946:GLU:HG2	2.00	0.43
1:A:772:VAL:HG13	1:A:788:LEU:HD23	2.01	0.42
1:A:839:ARG:HA	1:A:839:ARG:HD3	1.93	0.42
1:A:514:VAL:HG11	1:A:519:TYR:CD2	2.54	0.42
1:A:572:LEU:HB3	1:A:607:LEU:HD12	2.00	0.42
1:A:1135:ILE:O	1:A:1136:ILE:HB	2.20	0.42
1:A:888:ILE:HA	1:A:889:PRO:HD3	1.90	0.42
1:A:752:LYS:HB3	1:A:779:SER:HB3	2.01	0.41
1:A:1047:ARG:HD2	1:A:1047:ARG:O	2.20	0.41
1:A:1130:GLU:CB	1:A:1141:VAL:HG11	2.51	0.41
1:A:687:MET:O	1:A:697:ARG:NH1	2.53	0.41
1:A:530:ASN:OD1	1:A:591:ASN:ND2	2.54	0.40
1:A:1139:ILE:HG21	1:A:1144:TYR:HE2	1.86	0.40

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	725/751 (96%)	704 (97%)	20 (3%)	1 (0%)	51 62

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1136	ILE

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	642/685 (94%)	615 (96%)	27 (4%)	30 37

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	472	ARG
1	A	507	CYS
1	A	514	VAL
1	A	528	LEU
1	A	544	GLU
1	A	555	ASP
1	A	618	GLN
1	A	646	ILE
1	A	684	SER
1	A	737	ARG
1	A	766	CYS
1	A	828	ARG
1	A	840	SER
1	A	880	ILE
1	A	882	THR
1	A	924	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	954	LYS
1	A	965	GLU
1	A	967	LEU
1	A	969	ASP
1	A	1040	LEU
1	A	1047	ARG
1	A	1056	SER
1	A	1064	LYS
1	A	1117	ILE
1	A	1154	SER
1	A	1168	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	A	790	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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

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, 95^{th} 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	731/751 (97%)	-0.01	24 (3%) 46 57	12, 29, 62, 82	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	560	ASN	4.8
1	A	1122	LEU	4.6
1	A	1130	GLU	4.1
1	A	1111	LYS	3.7
1	A	644	TRP	3.6
1	A	882	THR	3.5
1	A	1117	ILE	3.5
1	A	470	ARG	3.5
1	A	471	GLU	3.2
1	A	1137	ALA	2.9
1	A	555	ASP	2.9
1	A	620	GLY	2.9
1	A	1115	GLU	2.8
1	A	637	GLU	2.7
1	A	622	GLN	2.5
1	A	618	GLN	2.4
1	A	474	ILE	2.4
1	A	1110	ASN	2.4
1	A	566	GLY	2.4
1	A	624	TYR	2.3
1	A	621	SER	2.3
1	A	634	GLN	2.1
1	A	712	HIS	2.1
1	A	1134	GLU	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 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.