



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 31, 2021 – 08:01 pm BST

PDB ID : 7ADA
Title : Crystal structure of helicase Pif1 from *Thermus oshimai* mutant Q164C-E409C
Authors : Dai, Y.X.; Chen, W.F.; Teng, F.Y.; Liu, N.N.; Hou, X.M.; Dou, S.X.; Rety, S.; Xi, X.G.
Deposited on : 2020-09-14
Resolution : 3.34 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 : 2.18
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) : 2.18

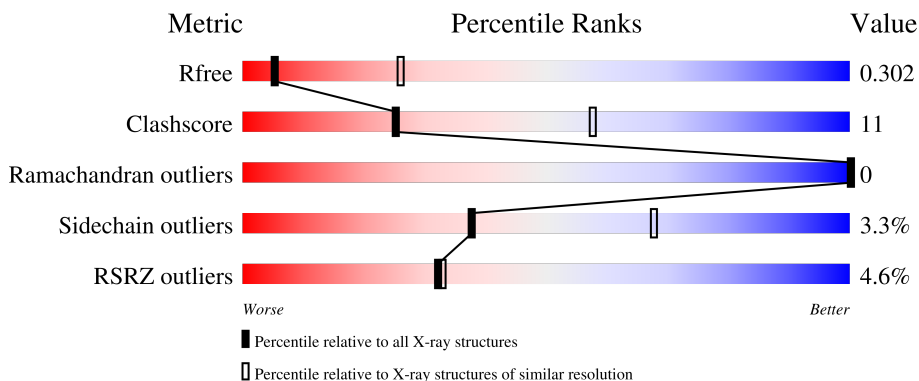
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

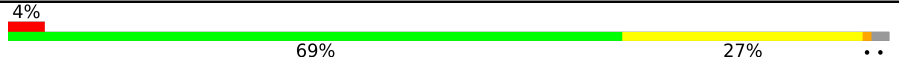
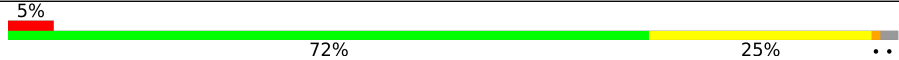
The reported resolution of this entry is 3.34 Å.

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	1060 (3.38-3.30)
Clashscore	141614	1111 (3.38-3.30)
Ramachandran outliers	138981	1090 (3.38-3.30)
Sidechain outliers	138945	1089 (3.38-3.30)
RSRZ outliers	127900	1028 (3.38-3.30)

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 $\leq 5\%$. 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	446	
1	B	446	

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 6996 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 PIF1 helicase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	436	3498	2238	640	615	5	0	0	0
1	B	436	3498	2238	640	615	5	0	0	0

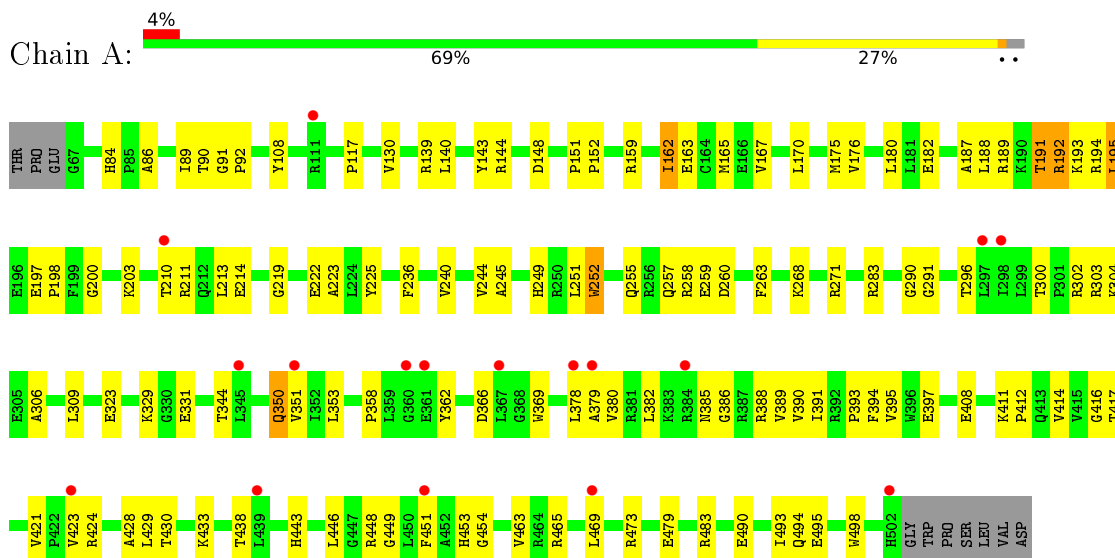
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	64	THR	-	expression tag	UNP K7RJ88
A	106	LYS	GLU	conflict	UNP K7RJ88
A	162	ILE	MET	conflict	UNP K7RJ88
A	164	CYS	GLN	engineered mutation	UNP K7RJ88
A	409	CYS	GLU	engineered mutation	UNP K7RJ88
A	456	LEU	PRO	conflict	UNP K7RJ88
A	508	VAL	-	expression tag	UNP K7RJ88
A	509	ASP	-	expression tag	UNP K7RJ88
B	64	THR	-	expression tag	UNP K7RJ88
B	106	LYS	GLU	conflict	UNP K7RJ88
B	162	ILE	MET	conflict	UNP K7RJ88
B	164	CYS	GLN	engineered mutation	UNP K7RJ88
B	409	CYS	GLU	engineered mutation	UNP K7RJ88
B	456	LEU	PRO	conflict	UNP K7RJ88
B	508	VAL	-	expression tag	UNP K7RJ88
B	509	ASP	-	expression tag	UNP K7RJ88

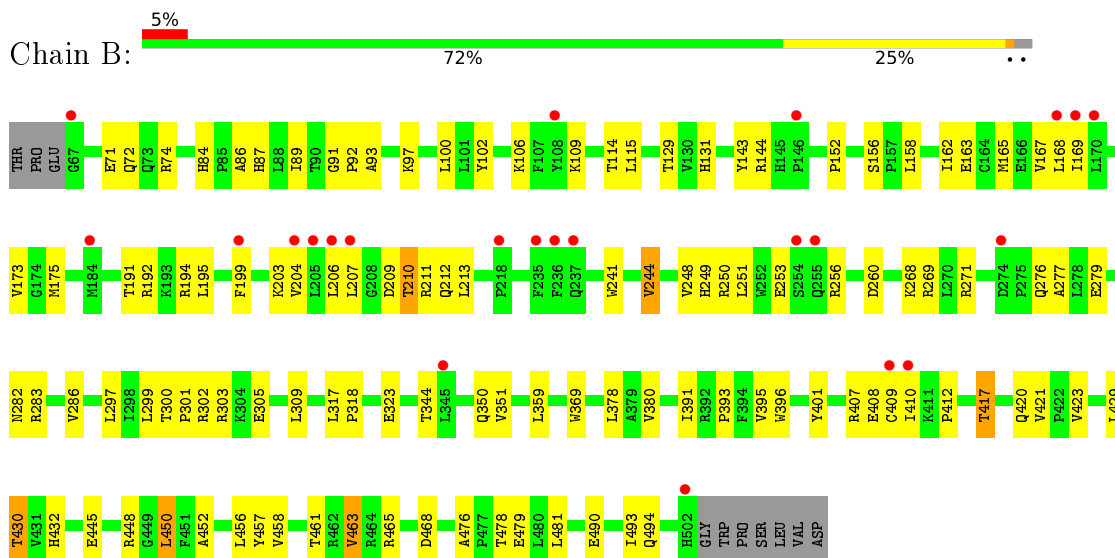
3 Residue-property plots [i](#)

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: PIF1 helicase



- Molecule 1: PIF1 helicase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	74.07Å 58.68Å 118.16Å 90.00° 92.21° 90.00°	Depositor
Resolution (Å)	52.55 – 3.34 52.55 – 3.34	Depositor EDS
% Data completeness (in resolution range)	99.6 (52.55-3.34) 99.8 (52.55-3.34)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.47 (at 3.33Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.245 , 0.302 0.244 , 0.302	Depositor DCC
R_{free} test set	754 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	137.3	Xtrriage
Anisotropy	0.553	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 110.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.021 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6996	wwPDB-VP
Average B, all atoms (Å ²)	166.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.57% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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 [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	0.27	0/3581	0.52	0/4859
1	B	0.28	0/3581	0.51	0/4859
All	All	0.27	0/7162	0.52	0/9718

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	4
1	B	0	1
All	All	0	5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	108	TYR	Peptide
1	A	283	ARG	Peptide
1	A	408	GLU	Peptide
1	A	494	GLN	Peptide
1	B	283	ARG	Peptide

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	3498	0	3563	84	0
1	B	3498	0	3563	68	0
All	All	6996	0	7126	151	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 11.

The worst 5 of 151 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:451:PHE:HB3	1:A:479:GLU:HG2	1.49	0.94
1:A:151:PRO:HB3	1:A:191:THR:HG23	1.67	0.77
1:A:130:VAL:HG11	1:A:180:LEU:HD21	1.67	0.76
1:A:92:PRO:HD3	1:A:252:TRP:HD1	1.50	0.76
1:B:106:LYS:HA	1:B:109:LYS:HG3	1.70	0.74

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	434/446 (97%)	426 (98%)	8 (2%)	0	100	100
1	B	434/446 (97%)	425 (98%)	9 (2%)	0	100	100
All	All	868/892 (97%)	851 (98%)	17 (2%)	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	364/373 (98%)	353 (97%)	11 (3%)	41 70
1	B	364/373 (98%)	351 (96%)	13 (4%)	35 65
All	All	728/746 (98%)	704 (97%)	24 (3%)	38 68

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	244	VAL
1	B	408	GLU
1	B	378	LEU
1	B	417	THR
1	A	252	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	249	HIS
1	A	385	ASN
1	B	276	GLN

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	436/446 (97%)	0.17	17 (3%) 39 38	99, 159, 218, 281	0
1	B	436/446 (97%)	0.29	23 (5%) 26 27	104, 161, 248, 292	0
All	All	872/892 (97%)	0.23	40 (4%) 32 33	99, 161, 236, 292	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	204	VAL	5.1
1	B	205	LEU	4.8
1	B	255	GLN	4.8
1	A	297	LEU	4.4
1	B	502	HIS	3.9

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.