Full wwPDB X-ray Structure Validation Report ⓘ

Jan 6, 2025 – 01:04 PM JST

PDB ID	:	8YTY
Title	:	The M12+P185V variant of Kubu-PETase from Kutzneria buriramensis
Authors	:	Park, J.; Seo, H.; Hong, H.; Kim, K.-J.
Deposited on	:	2024-03-26
Resolution	:	1.15 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

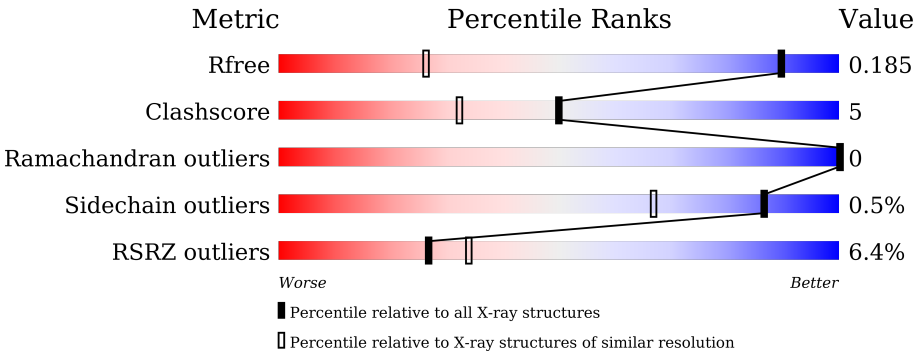
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 1.15 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	1095 (1.16-1.12)
Clashscore	180529	1248 (1.16-1.12)
Ramachandran outliers	177936	1224 (1.16-1.12)
Sidechain outliers	177891	1224 (1.16-1.12)
RSRZ outliers	164620	1095 (1.16-1.12)

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	260	6% 83% 13% .

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2322 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 Dienelactone hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	251	1955	1243	340	363	9	0	18	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	MET	-	initiating methionine	UNP A0A3E0H050
A	95	ARG	THR	variant	UNP A0A3E0H050
A	119	ASN	THR	variant	UNP A0A3E0H050
A	127	SER	GLN	variant	UNP A0A3E0H050
A	131	GLN	GLU	variant	UNP A0A3E0H050
A	173	CYS	ALA	variant	UNP A0A3E0H050
A	184	ILE	VAL	variant	UNP A0A3E0H050
A	185	VAL	PRO	variant	UNP A0A3E0H050
A	190	HIS	ASP	variant	UNP A0A3E0H050
A	197	CYS	LYS	variant	UNP A0A3E0H050
A	236	CYS	ALA	variant	UNP A0A3E0H050
A	239	SER	ASP	variant	UNP A0A3E0H050
A	279	SER	ALA	variant	UNP A0A3E0H050
A	281	CYS	SER	variant	UNP A0A3E0H050
A	289	LEU	-	expression tag	UNP A0A3E0H050
A	290	GLU	-	expression tag	UNP A0A3E0H050
A	291	HIS	-	expression tag	UNP A0A3E0H050
A	292	HIS	-	expression tag	UNP A0A3E0H050
A	293	HIS	-	expression tag	UNP A0A3E0H050
A	294	HIS	-	expression tag	UNP A0A3E0H050
A	295	HIS	-	expression tag	UNP A0A3E0H050
A	296	HIS	-	expression tag	UNP A0A3E0H050

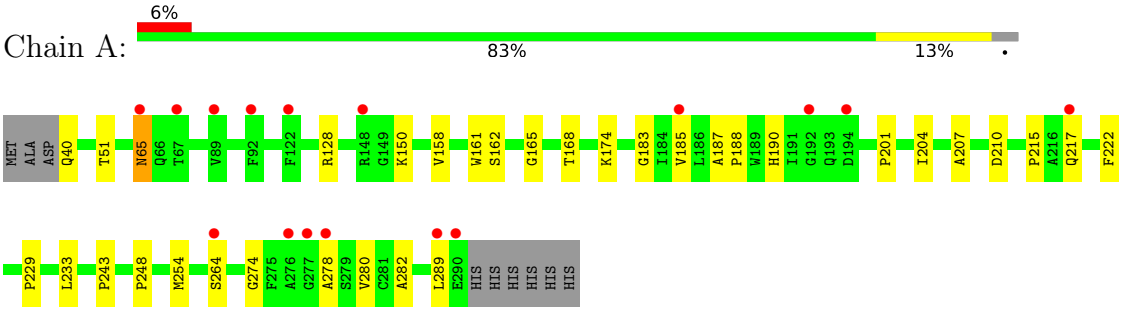
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	367	Total O 367 367	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 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: Dienelactone hydrolase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	64.96Å 81.64Å 83.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.50 – 1.15 32.50 – 1.15	Depositor EDS
% Data completeness (in resolution range)	97.9 (32.50-1.15) 98.2 (32.50-1.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.53 (at 1.14Å)	Xtriage
Refinement program	REFMAC 5.8.0411	Depositor
R, R_{free}	0.160 , 0.178 0.168 , 0.185	Depositor DCC
R_{free} test set	4030 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	8.9	Xtriage
Anisotropy	0.123	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 33.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.009 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2322	wwPDB-VP
Average B, all atoms (Å^2)	11.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.60% 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	0.77	4/2017 (0.2%)	0.87	4/2760 (0.1%)

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	5

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	215[A]	PRO	N-CA	5.37	1.56	1.47
1	A	215[B]	PRO	N-CA	5.37	1.56	1.47
1	A	215[A]	PRO	C-O	5.04	1.33	1.23
1	A	215[B]	PRO	C-O	5.04	1.33	1.23

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	215[A]	PRO	CA-N-CD	-7.32	101.25	111.50
1	A	215[B]	PRO	CA-N-CD	-7.32	101.25	111.50
1	A	278	ALA	N-CA-CB	-6.24	101.37	110.10
1	A	40	GLN	N-CA-CB	5.20	119.96	110.60

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	128	ARG	Sidechain
1	A	264[B]	SER	Mainchain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	274[B]	GLY	Mainchain
1	A	274[C]	GLY	Mainchain
1	A	65[B]	ASN	Mainchain

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	1955	0	1896	20	1
2	A	367	0	0	3	2
All	All	2322	0	1896	20	3

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 (20) 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:217:GLN:O	2:A:301:HOH:O	1.95	0.83
1:A:174[B]:LYS:HD2	2:A:435:HOH:O	1.91	0.70
1:A:185[B]:VAL:CG1	1:A:188:PRO:HB3	2.24	0.68
1:A:233[B]:LEU:HD11	1:A:280:VAL:HG13	1.79	0.64
1:A:150[A]:LYS:NZ	2:A:302:HOH:O	2.33	0.60
1:A:185[B]:VAL:HG12	1:A:188:PRO:HB3	1.86	0.56
1:A:233[B]:LEU:CD1	1:A:280:VAL:HG13	2.38	0.54
1:A:51:THR:HB	1:A:248[A]:PRO:HB3	1.91	0.52
1:A:161:TRP:CD1	1:A:243:PRO:HG2	2.51	0.46
1:A:168:THR:HB	1:A:185[A]:VAL:HG12	1.98	0.46
1:A:233[B]:LEU:HD12	1:A:282:ALA:O	2.17	0.45
1:A:165:GLY:CA	1:A:185[B]:VAL:HG13	2.47	0.44
1:A:162:SER:HA	1:A:187:ALA:O	2.18	0.44
1:A:174[B]:LYS:HE3	1:A:174[B]:LYS:HB2	1.63	0.43
1:A:185[A]:VAL:CG2	1:A:204:ILE:HG12	2.49	0.43
1:A:190:HIS:HB3	1:A:222:PHE:CZ	2.54	0.42
1:A:158:VAL:O	1:A:183:GLY:HA2	2.18	0.42
1:A:201:PRO:HA	1:A:229:PRO:O	2.21	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:185[B]:VAL:HG11	1:A:188:PRO:HB3	2.03	0.41
1:A:207:ALA:HB3	1:A:210:ASP:HB2	2.03	0.40

All (3) 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)
2:A:313:HOH:O	2:A:313:HOH:O[2_555]	1.38	0.82
2:A:317:HOH:O	2:A:505:HOH:O[8_544]	1.90	0.30
1:A:65[B]:ASN:O	1:A:289:LEU:CD2[6_445]	2.12	0.08

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	267/260 (103%)	261 (98%)	6 (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	203/194 (105%)	201 (99%)	2 (1%)	73 40

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	254[A]	MET
1	A	254[B]	MET

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

Mol	Chain	Res	Type
1	A	119	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)

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

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	251/260 (96%)	0.34	16 (6%) 27 34	3, 8, 14, 36	18 (7%)

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	278	ALA	5.3
1	A	194	ASP	5.1
1	A	289	LEU	4.5
1	A	276	ALA	4.3
1	A	277	GLY	3.4
1	A	290	GLU	2.9
1	A	185[A]	VAL	2.9
1	A	122	PHE	2.8
1	A	192	GLY	2.8
1	A	92	PHE	2.8
1	A	65[A]	ASN	2.7
1	A	264[A]	SER	2.7
1	A	67	THR	2.6
1	A	148[A]	ARG	2.3
1	A	89[A]	VAL	2.1
1	A	217	GLN	2.1

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.