



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 16, 2024 – 06:09 PM JST

PDB ID : 8H83  
Title : Crystal structure of a IsPETase variant V22 from Ideonella sakaiensis  
Authors : Wei, H.L.; Gao, S.F.; Li, Q.; Han, X.; Gao, J.; Liu, W.D.  
Deposited on : 2022-10-21  
Resolution : 1.92 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](mailto: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) : 1.13  
EDS : 2.36  
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) : 2.36

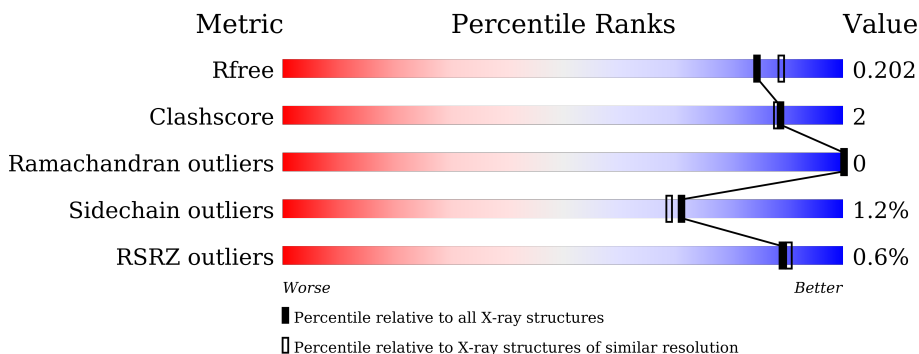
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.92 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	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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  $\geq 3$ , 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	330	 74% 5% 21%
1	B	330	 75% 5% 21%

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 4285 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 Poly(ethylene terephthalate) hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	261	1919	1186	344	378	11	0	0	0
1	B	260	1913	1183	343	376	11	0	0	0

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-31	MET	-	initiating methionine	UNP A0A0K8P6T7
A	-30	LYS	-	expression tag	UNP A0A0K8P6T7
A	-29	TYR	-	expression tag	UNP A0A0K8P6T7
A	-28	LEU	-	expression tag	UNP A0A0K8P6T7
A	-27	LEU	-	expression tag	UNP A0A0K8P6T7
A	-26	PRO	-	expression tag	UNP A0A0K8P6T7
A	-25	THR	-	expression tag	UNP A0A0K8P6T7
A	-24	ALA	-	expression tag	UNP A0A0K8P6T7
A	-23	ALA	-	expression tag	UNP A0A0K8P6T7
A	-22	ALA	-	expression tag	UNP A0A0K8P6T7
A	-21	GLY	-	expression tag	UNP A0A0K8P6T7
A	-20	LEU	-	expression tag	UNP A0A0K8P6T7
A	-19	LEU	-	expression tag	UNP A0A0K8P6T7
A	-18	LEU	-	expression tag	UNP A0A0K8P6T7
A	-17	LEU	-	expression tag	UNP A0A0K8P6T7
A	-16	ALA	-	expression tag	UNP A0A0K8P6T7
A	-15	ALA	-	expression tag	UNP A0A0K8P6T7
A	-14	GLN	-	expression tag	UNP A0A0K8P6T7
A	-13	PRO	-	expression tag	UNP A0A0K8P6T7
A	-12	THR	-	expression tag	UNP A0A0K8P6T7
A	-11	MET	-	expression tag	UNP A0A0K8P6T7
A	-10	ALA	-	expression tag	UNP A0A0K8P6T7
A	-9	MET	-	expression tag	UNP A0A0K8P6T7
A	-8	ASP	-	expression tag	UNP A0A0K8P6T7
A	-7	ILE	-	expression tag	UNP A0A0K8P6T7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	GLY	-	expression tag	UNP A0A0K8P6T7
A	-5	ILE	-	expression tag	UNP A0A0K8P6T7
A	-4	ASN	-	expression tag	UNP A0A0K8P6T7
A	-3	SER	-	expression tag	UNP A0A0K8P6T7
A	-2	ASP	-	expression tag	UNP A0A0K8P6T7
A	-1	PRO	-	expression tag	UNP A0A0K8P6T7
A	0	ASN	-	expression tag	UNP A0A0K8P6T7
A	1	SER	-	expression tag	UNP A0A0K8P6T7
A	14	GLU	VAL	variant	UNP A0A0K8P6T7
A	18	PRO	LEU	variant	UNP A0A0K8P6T7
A	84	LEU	VAL	variant	UNP A0A0K8P6T7
A	201	ILE	PHE	variant	UNP A0A0K8P6T7
A	204	GLN	GLU	variant	UNP A0A0K8P6T7
A	229	TYR	PHE	variant	UNP A0A0K8P6T7
A	233	LYS	ASN	variant	UNP A0A0K8P6T7
A	280	GLU	ARG	variant	UNP A0A0K8P6T7
A	283	ARG	ASP	variant	UNP A0A0K8P6T7
A	291	LEU	-	expression tag	UNP A0A0K8P6T7
A	292	GLU	-	expression tag	UNP A0A0K8P6T7
A	293	HIS	-	expression tag	UNP A0A0K8P6T7
A	294	HIS	-	expression tag	UNP A0A0K8P6T7
A	295	HIS	-	expression tag	UNP A0A0K8P6T7
A	296	HIS	-	expression tag	UNP A0A0K8P6T7
A	297	HIS	-	expression tag	UNP A0A0K8P6T7
A	298	HIS	-	expression tag	UNP A0A0K8P6T7
B	-31	MET	-	initiating methionine	UNP A0A0K8P6T7
B	-30	LYS	-	expression tag	UNP A0A0K8P6T7
B	-29	TYR	-	expression tag	UNP A0A0K8P6T7
B	-28	LEU	-	expression tag	UNP A0A0K8P6T7
B	-27	LEU	-	expression tag	UNP A0A0K8P6T7
B	-26	PRO	-	expression tag	UNP A0A0K8P6T7
B	-25	THR	-	expression tag	UNP A0A0K8P6T7
B	-24	ALA	-	expression tag	UNP A0A0K8P6T7
B	-23	ALA	-	expression tag	UNP A0A0K8P6T7
B	-22	ALA	-	expression tag	UNP A0A0K8P6T7
B	-21	GLY	-	expression tag	UNP A0A0K8P6T7
B	-20	LEU	-	expression tag	UNP A0A0K8P6T7
B	-19	LEU	-	expression tag	UNP A0A0K8P6T7
B	-18	LEU	-	expression tag	UNP A0A0K8P6T7
B	-17	LEU	-	expression tag	UNP A0A0K8P6T7
B	-16	ALA	-	expression tag	UNP A0A0K8P6T7
B	-15	ALA	-	expression tag	UNP A0A0K8P6T7

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	GLN	-	expression tag	UNP A0A0K8P6T7
B	-13	PRO	-	expression tag	UNP A0A0K8P6T7
B	-12	THR	-	expression tag	UNP A0A0K8P6T7
B	-11	MET	-	expression tag	UNP A0A0K8P6T7
B	-10	ALA	-	expression tag	UNP A0A0K8P6T7
B	-9	MET	-	expression tag	UNP A0A0K8P6T7
B	-8	ASP	-	expression tag	UNP A0A0K8P6T7
B	-7	ILE	-	expression tag	UNP A0A0K8P6T7
B	-6	GLY	-	expression tag	UNP A0A0K8P6T7
B	-5	ILE	-	expression tag	UNP A0A0K8P6T7
B	-4	ASN	-	expression tag	UNP A0A0K8P6T7
B	-3	SER	-	expression tag	UNP A0A0K8P6T7
B	-2	ASP	-	expression tag	UNP A0A0K8P6T7
B	-1	PRO	-	expression tag	UNP A0A0K8P6T7
B	0	ASN	-	expression tag	UNP A0A0K8P6T7
B	1	SER	-	expression tag	UNP A0A0K8P6T7
B	14	GLU	VAL	variant	UNP A0A0K8P6T7
B	18	PRO	LEU	variant	UNP A0A0K8P6T7
B	84	LEU	VAL	variant	UNP A0A0K8P6T7
B	201	ILE	PHE	variant	UNP A0A0K8P6T7
B	204	GLN	GLU	variant	UNP A0A0K8P6T7
B	229	TYR	PHE	variant	UNP A0A0K8P6T7
B	233	LYS	ASN	variant	UNP A0A0K8P6T7
B	280	GLU	ARG	variant	UNP A0A0K8P6T7
B	283	ARG	ASP	variant	UNP A0A0K8P6T7
B	291	LEU	-	expression tag	UNP A0A0K8P6T7
B	292	GLU	-	expression tag	UNP A0A0K8P6T7
B	293	HIS	-	expression tag	UNP A0A0K8P6T7
B	294	HIS	-	expression tag	UNP A0A0K8P6T7
B	295	HIS	-	expression tag	UNP A0A0K8P6T7
B	296	HIS	-	expression tag	UNP A0A0K8P6T7
B	297	HIS	-	expression tag	UNP A0A0K8P6T7
B	298	HIS	-	expression tag	UNP A0A0K8P6T7

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	227	Total O 227 227	0	0
2	B	226	Total O 226 226	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: Poly(ethylene terephthalate) hydrolase

Chain A:  74% 5% 21%

MET LYS TYR LEU LEU PRO THR ALA ALA ALA GLY LEU LEU LEU LEU ALA ALA ALA GLN PRO THR MET MET MET ASP ILE GLY ILE ASN SER ASP PRO ASN SER PHE PRO ARG ALA ALA ARG ARG LEU MET MET GLN ALA ALA ALA LEU LEU LEU LEU PRO MET MET VAL SER ALA ALA ALA THR ALA GLN

THR R30 P36 P49 R53 Y70 L84 T88 T116 R123 Q127 S141 Y146 S160 A183 D220 L230 K253 S290 LEU GLU HIS HIS HIS HIS HIS

- Molecule 1: Poly(ethylene terephthalate) hydrolase

Chain B:  % 75% •• 21%

MET LYS TYR LEU LEU PRO THR ALA ALA ALA GLY LEU LEU LEU LEU ALA ALA ALA GLN PRO THR MET MET MET ASP ILE GLY ILE ASN SER ASP PRO ASN SER PHE PRO ARG ALA ALA ARG ARG LEU MET MET GLN ALA ALA ALA LEU LEU LEU LEU PRO MET MET VAL SER ALA ALA ALA THR ALA GLN

THR R30 R59 T88 T113 T116 R123 S141 Y146 S160 A183 L216 P217 D220 R283 C289 SER LEU GLU HIS HIS HIS HIS HIS HIS HIS

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.23Å 80.66Å 134.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.89 – 1.92 47.89 – 1.92	Depositor EDS
% Data completeness (in resolution range)	96.7 (47.89-1.92) 87.2 (47.89-1.92)	Depositor EDS
$R_{merge}$	0.22	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.91 (at 1.92Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.192 , 0.208 0.186 , 0.202	Depositor DCC
$R_{free}$ test set	2099 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.9	Xtrriage
Anisotropy	0.065	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 49.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4285	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 47.49 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.8472e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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	0.37	0/1963	0.59	0/2673
1	B	0.40	0/1957	0.60	0/2665
All	All	0.39	0/3920	0.60	0/5338

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 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	1919	0	1860	9	0
1	B	1913	0	1855	6	0
2	A	227	0	0	0	0
2	B	226	0	0	0	0
All	All	4285	0	3715	14	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 2.

The worst 5 of 14 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:160:SER:HA	1:A:183:ALA:O	2.12	0.50

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:216:LEU:N	1:B:217:PRO:HD2	2.26	0.50
1:B:88:THR:HA	1:B:116:THR:HB	1.95	0.48
1:A:230:LEU:HD22	1:A:253:LYS:HG3	1.96	0.48
1:B:160:SER:HA	1:B:183:ALA:O	2.14	0.47

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	259/330 (78%)	255 (98%)	4 (2%)	0	100	100
1	B	258/330 (78%)	254 (98%)	4 (2%)	0	100	100
All	All	517/660 (78%)	509 (98%)	8 (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	207/258 (80%)	205 (99%)	2 (1%)	76	75
1	B	206/258 (80%)	203 (98%)	3 (2%)	65	61
All	All	413/516 (80%)	408 (99%)	5 (1%)	71	69

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	123	ARG
1	A	160	SER
1	B	59	ARG
1	B	123	ARG
1	B	146	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	261/330 (79%)	-0.23	1 (0%) 92 93	15, 22, 31, 43	0
1	B	260/330 (78%)	-0.16	2 (0%) 86 87	16, 22, 32, 41	0
All	All	521/660 (78%)	-0.19	3 (0%) 89 90	15, 22, 32, 43	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	146	TYR	3.5
1	A	290	SER	3.2
1	B	283	ARG	2.3

### 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.