wwPDB X-ray Structure Validation Summary Report (i)

Mar 8, 2026 – 06:55 AM UTC

PDB ID	:	8H5O / pdb_00008h5o
Title	:	Crystal structure of PETase S121E/P181V/D186H/N233C/S242T/N246D/S282C mutant from Ideonella sakaiensis
Authors	:	Lee, S.H.; Seo, H.; Kim, K.-J.
Deposited on	:	2022-10-13
Resolution	:	2.10 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 (i) 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 (1)) were used in the production of this report:

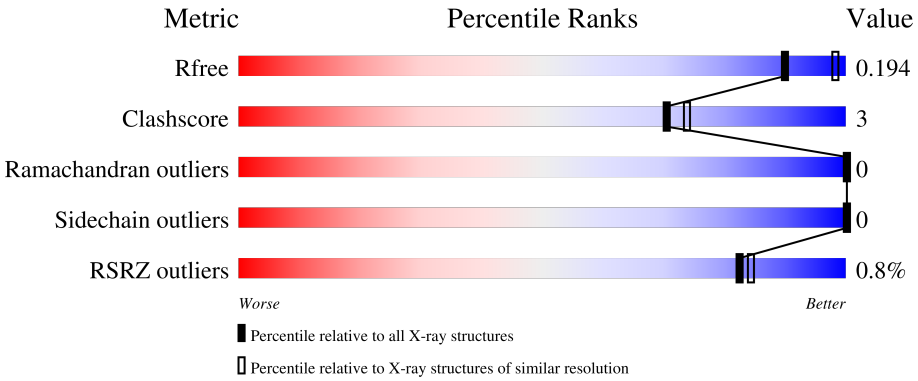
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 2.10 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}	180053	6658 (2.10-2.10)
Clashscore	190562	7164 (2.10-2.10)
Ramachandran outliers	187476	7099 (2.10-2.10)
Sidechain outliers	187428	7100 (2.10-2.10)
RSRZ outliers	180081	6662 (2.10-2.10)

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	300	% 79% 8% 13%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2096 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	262	1948	1201	350	382	15	0	3	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	13	MET	-	initiating methionine	UNP A0A0K8P6T7
A	14	GLY	-	expression tag	UNP A0A0K8P6T7
A	15	SER	-	expression tag	UNP A0A0K8P6T7
A	16	SER	-	expression tag	UNP A0A0K8P6T7
A	17	HIS	-	expression tag	UNP A0A0K8P6T7
A	18	HIS	-	expression tag	UNP A0A0K8P6T7
A	19	HIS	-	expression tag	UNP A0A0K8P6T7
A	20	HIS	-	expression tag	UNP A0A0K8P6T7
A	21	HIS	-	expression tag	UNP A0A0K8P6T7
A	22	HIS	-	expression tag	UNP A0A0K8P6T7
A	23	SER	-	expression tag	UNP A0A0K8P6T7
A	24	SER	-	expression tag	UNP A0A0K8P6T7
A	25	GLY	-	expression tag	UNP A0A0K8P6T7
A	26	LEU	-	expression tag	UNP A0A0K8P6T7
A	27	VAL	-	expression tag	UNP A0A0K8P6T7
A	28	PRO	-	expression tag	UNP A0A0K8P6T7
A	29	ARG	-	expression tag	UNP A0A0K8P6T7
A	30	GLY	-	expression tag	UNP A0A0K8P6T7
A	31	SER	-	expression tag	UNP A0A0K8P6T7
A	32	HIS	-	expression tag	UNP A0A0K8P6T7
A	33	MET	-	expression tag	UNP A0A0K8P6T7
A	121	GLU	SER	engineered mutation	UNP A0A0K8P6T7
A	181	VAL	PRO	engineered mutation	UNP A0A0K8P6T7
A	186	HIS	ASP	engineered mutation	UNP A0A0K8P6T7
A	233	CYS	ASN	engineered mutation	UNP A0A0K8P6T7
A	242	THR	SER	engineered mutation	UNP A0A0K8P6T7
A	246	ASP	ASN	engineered mutation	UNP A0A0K8P6T7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	282	CYS	SER	engineered mutation	UNP A0A0K8P6T7
A	291	LEU	-	expression tag	UNP A0A0K8P6T7
A	292	GLU	-	expression tag	UNP A0A0K8P6T7
A	293	ASP	-	expression tag	UNP A0A0K8P6T7
A	294	PRO	-	expression tag	UNP A0A0K8P6T7
A	295	ALA	-	expression tag	UNP A0A0K8P6T7
A	296	ALA	-	expression tag	UNP A0A0K8P6T7
A	297	ASN	-	expression tag	UNP A0A0K8P6T7
A	298	LYS	-	expression tag	UNP A0A0K8P6T7
A	299	ALA	-	expression tag	UNP A0A0K8P6T7
A	300	ARG	-	expression tag	UNP A0A0K8P6T7
A	301	LYS	-	expression tag	UNP A0A0K8P6T7
A	302	GLU	-	expression tag	UNP A0A0K8P6T7
A	303	ALA	-	expression tag	UNP A0A0K8P6T7
A	304	GLU	-	expression tag	UNP A0A0K8P6T7
A	305	LEU	-	expression tag	UNP A0A0K8P6T7
A	306	ALA	-	expression tag	UNP A0A0K8P6T7
A	307	ALA	-	expression tag	UNP A0A0K8P6T7
A	308	ALA	-	expression tag	UNP A0A0K8P6T7
A	309	THR	-	expression tag	UNP A0A0K8P6T7
A	310	ALA	-	expression tag	UNP A0A0K8P6T7
A	311	GLU	-	expression tag	UNP A0A0K8P6T7
A	312	GLN	-	expression tag	UNP A0A0K8P6T7

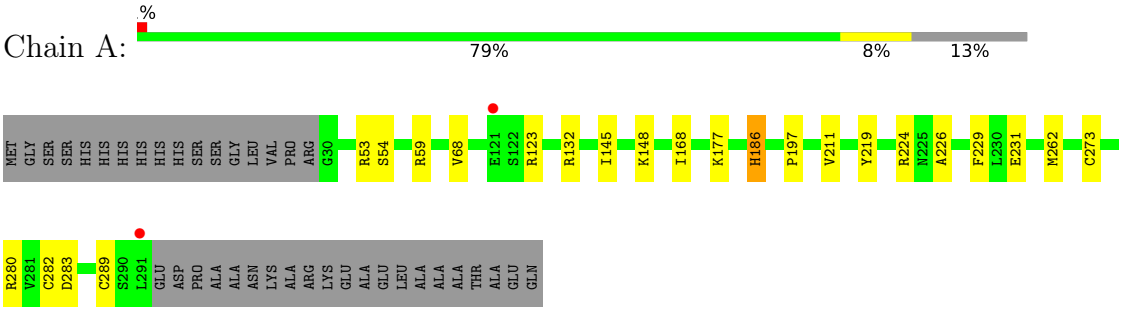
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	148	Total O 148 148	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



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	50.86Å 52.62Å 95.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.01 – 2.10 29.01 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.2 (29.01-2.10) 99.5 (29.01-2.10)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	15.04 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0352	Depositor
R, R_{free}	0.141 , 0.186 0.155 , 0.194	Depositor DCC
R_{free} test set	756 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å^2)	22.6	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 52.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.027 for k,h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2096	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.36% 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	0.70	1/1992 (0.1%)	1.00	1/2710 (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	5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	186	HIS	CE1-NE2	6.33	1.38	1.32

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	283	ASP	CA-CB-CG	5.03	117.63	112.60

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	123	ARG	Sidechain
1	A	224	ARG	Sidechain
1	A	280	ARG	Sidechain
1	A	53	ARG	Sidechain
1	A	59	ARG	Sidechain

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	1948	0	1874	12	0
2	A	148	0	0	1	0
All	All	2096	0	1874	12	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 3.

The worst 5 of 12 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:273:CYS:SG	1:A:289[B]:CYS:HB3	2.21	0.80
1:A:273:CYS:HG	1:A:289[B]:CYS:HB3	1.55	0.69
1:A:168:ILE:HD11	1:A:186:HIS:CE1	2.39	0.58
1:A:211:VAL:HG11	1:A:231:GLU:OE2	2.05	0.55
1:A:177:LYS:HB2	1:A:262:MET:HG2	1.94	0.48

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	263/300 (88%)	257 (98%)	6 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	211/236 (89%)	211 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	32	HIS
1	A	247	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 oligosaccharides 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 (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	262/300 (87%)	-0.34	2 (0%)	82	84	13, 22, 34, 61	3 (1%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	291	LEU	4.3
1	A	121	GLU	2.2

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 oligosaccharides 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.