Full wwPDB X-ray Structure Validation Report ⓘ

Apr 21, 2026 – 12:31 PM JST

PDB ID	:	9UM8 / pdb_00009um8
Title	:	CaPETaseM9 + P289E variant
Authors	:	Kim, K.; Ki, D.; Park, J.
Deposited on	:	2025-04-21
Resolution	:	1.22 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:

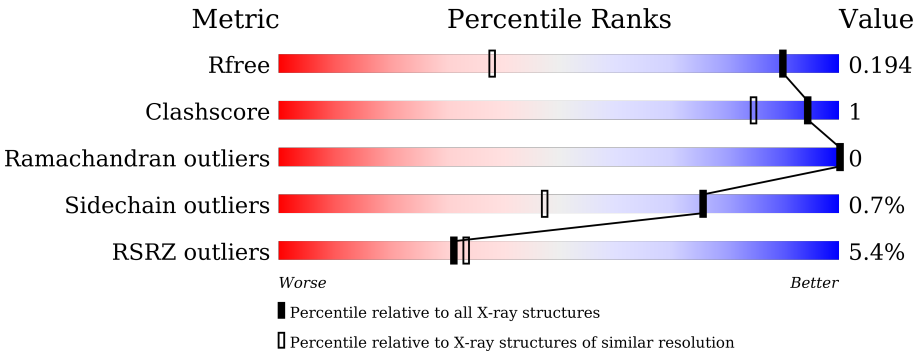
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 1.22 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	2002 (1.24-1.20)
Clashscore	190562	2061 (1.24-1.20)
Ramachandran outliers	187476	2009 (1.24-1.20)
Sidechain outliers	187428	2008 (1.24-1.20)
RSRZ outliers	180081	2000 (1.24-1.20)

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	263	5% 92% 6% •
1	B	263	6% 95% 5%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4273 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 Cutinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	260	Total	C	N	O	S	0	0	0
			1978	1256	341	371	10			
1	B	263	Total	C	N	O	S	0	0	0
			1998	1268	346	374	10			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	109	ALA	ASN	engineered mutation	UNP A0A1M7II12
A	129	THR	VAL	engineered mutation	UNP A0A1M7II12
A	155	ARG	ALA	engineered mutation	UNP A0A1M7II12
A	180	CYS	LEU	engineered mutation	UNP A0A1M7II12
A	196	THR	GLY	engineered mutation	UNP A0A1M7II12
A	198	LYS	ARG	engineered mutation	UNP A0A1M7II12
A	202	CYS	ALA	engineered mutation	UNP A0A1M7II12
A	242	CYS	ARG	engineered mutation	UNP A0A1M7II12
A	289	GLU	PRO	engineered mutation	UNP A0A1M7II12
A	291	CYS	SER	engineered mutation	UNP A0A1M7II12
A	300	LEU	-	expression tag	UNP A0A1M7II12
A	301	GLU	-	expression tag	UNP A0A1M7II12
A	302	HIS	-	expression tag	UNP A0A1M7II12
A	303	HIS	-	expression tag	UNP A0A1M7II12
B	109	ALA	ASN	engineered mutation	UNP A0A1M7II12
B	129	THR	VAL	engineered mutation	UNP A0A1M7II12
B	155	ARG	ALA	engineered mutation	UNP A0A1M7II12
B	180	CYS	LEU	engineered mutation	UNP A0A1M7II12
B	196	THR	GLY	engineered mutation	UNP A0A1M7II12
B	198	LYS	ARG	engineered mutation	UNP A0A1M7II12
B	202	CYS	ALA	engineered mutation	UNP A0A1M7II12
B	242	CYS	ARG	engineered mutation	UNP A0A1M7II12
B	289	GLU	PRO	engineered mutation	UNP A0A1M7II12
B	291	CYS	SER	engineered mutation	UNP A0A1M7II12
B	300	LEU	-	expression tag	UNP A0A1M7II12

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	301	GLU	-	expression tag	UNP A0A1M7II12
B	302	HIS	-	expression tag	UNP A0A1M7II12
B	303	HIS	-	expression tag	UNP A0A1M7II12

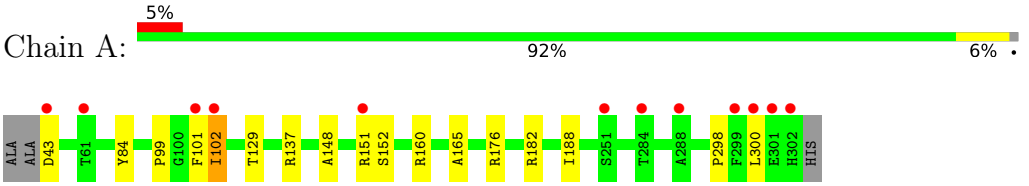
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	142	Total O 142 142	0	0
2	B	155	Total O 155 155	0	0

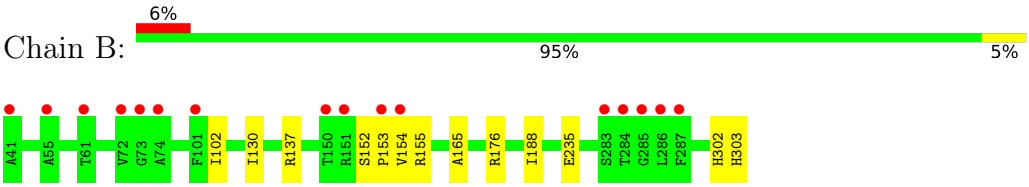
3 Residue-property plots (i)

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: Cutinase



• Molecule 1: Cutinase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	40.99Å 112.02Å 113.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.25 – 1.22 33.25 – 1.22	Depositor EDS
% Data completeness (in resolution range)	93.4 (33.25-1.22) 93.4 (33.25-1.22)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.67 (at 1.22Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.175 , 0.191 0.179 , 0.194	Depositor DCC
R_{free} test set	7306 reflections (4.70%)	wwPDB-VP
Wilson B-factor (Å^2)	11.9	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 27.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.019 for -h,l,k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4273	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% 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	1.03	0/2037	1.17	0/2787
1	B	1.00	0/2058	1.15	1/2816 (0.0%)
All	All	1.02	0/4095	1.16	1/5603 (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
1	B	0	3
All	All	0	8

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	B	153	PRO	N-CA-CB	-5.53	97.66	103.19

There are no chirality outliers.

All (8) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	137	ARG	Sidechain
1	A	151	ARG	Sidechain
1	A	160	ARG	Sidechain
1	A	176	ARG	Sidechain
1	A	182	ARG	Sidechain
1	B	137	ARG	Sidechain
1	B	155	ARG	Sidechain
1	B	176	ARG	Sidechain

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	1978	0	1895	6	0
1	B	1998	0	1912	5	0
2	A	142	0	0	1	0
2	B	155	0	0	0	0
All	All	4273	0	3807	10	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 1.

All (10) 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:298:PRO:HG2	1:A:300:LEU:HG	1.81	0.62
1:A:99:PRO:HB3	2:A:489:HOH:O	2.01	0.61
1:A:84:TYR:OH	1:B:303:HIS:HD2	1.84	0.59
1:B:235:GLU:OE1	1:B:302:HIS:HE1	1.88	0.57
1:B:102:ILE:HA	1:B:130:ILE:HG12	1.91	0.52
1:B:152:SER:OG	1:B:154:VAL:HG22	2.15	0.46
1:A:101:PHE:O	1:A:102:ILE:C	2.62	0.42
1:B:165:ALA:O	1:B:188:ILE:HA	2.20	0.41
1:A:165:ALA:O	1:A:188:ILE:HA	2.21	0.41
1:A:148:ALA:HA	1:A:152:SER:HB3	2.01	0.41

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	258/263 (98%)	251 (97%)	7 (3%)	0	100	100
1	B	261/263 (99%)	257 (98%)	4 (2%)	0	100	100
All	All	519/526 (99%)	508 (98%)	11 (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	207/208 (100%)	204 (99%)	3 (1%)	59	24
1	B	208/208 (100%)	208 (100%)	0	100	100
All	All	415/416 (100%)	412 (99%)	3 (1%)	76	47

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	43	ASP
1	A	102	ILE
1	A	129	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	69	GLN
1	A	270	ASN
1	B	81	GLN
1	B	90	GLN
1	B	276	GLN
1	B	302	HIS
1	B	303	HIS

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

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	260/263 (98%)	0.24	12 (4%)	37	40	7, 11, 19, 49	0
1	B	263/263 (100%)	0.29	16 (6%)	27	28	7, 12, 23, 40	0
All	All	523/526 (99%)	0.26	28 (5%)	31	33	7, 12, 21, 49	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	300	LEU	7.2
1	A	288	ALA	5.9
1	B	284	THR	5.5
1	B	41	ALA	5.4
1	B	286	LEU	4.2
1	A	301	GLU	4.2
1	A	102	ILE	4.2
1	A	43	ASP	4.1
1	B	283	SER	3.9
1	B	287	PHE	3.7
1	B	153	PRO	3.6
1	A	61	THR	3.3
1	A	284	THR	3.3
1	B	72	VAL	3.3
1	B	73	GLY	3.3
1	B	154	VAL	3.3
1	B	151	ARG	3.0
1	A	302	HIS	3.0
1	A	101	PHE	2.9
1	B	61	THR	2.6
1	A	299	PHE	2.6
1	A	151	ARG	2.5
1	B	55	ALA	2.5
1	B	101	PHE	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	150	THR	2.1
1	A	251	SER	2.0
1	B	74	ALA	2.0
1	B	285	GLY	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 oligosaccharides 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.