



wwPDB EM Validation Summary Report ⓘ

Jun 22, 2026 – 12:11 PM EDT

PDB ID : 9OZU / pdb_00009ozu
Title : MicroED structure of proteinase K from spray-frozen microcrystals
Authors : Summers, J.A.; Vlahakis, N.; Rodriguez, J.A.; Dahlberg, P.; Wakatsuki, S.
Deposited on : 2025-06-05
Resolution : 2.50 Å (reported)
Based on initial model : 2ID8

This is a wwPDB EM Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

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
Mogul : **NOT EXECUTED**
Buster-report : **NOT EXECUTED**
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

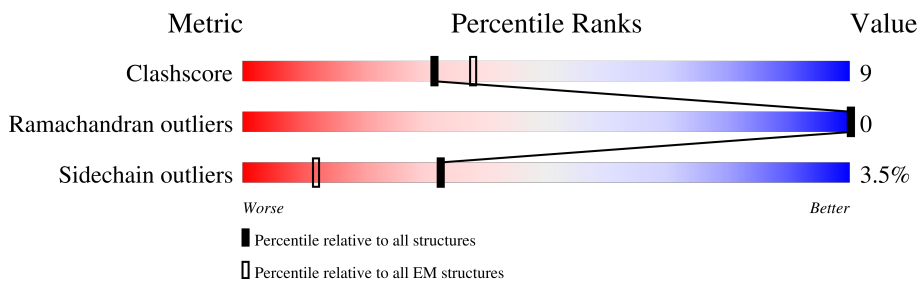
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON CRYSTALLOGRAPHY

The reported resolution of this entry is 2.50 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)	EM structures (#Entries)
Clashscore	229148	23984
Ramachandran outliers	224038	23583
Sidechain outliers	223484	23102

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	279	81% (green), 18% (yellow), 1% (orange), 0% (red), 0% (grey)
1	B	279	80% (green), 18% (yellow), 1% (orange), 0% (red), 0% (grey)

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 4108 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Proteinase K.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	279	Total	C	N	O	S	11	0
			2054	1255	361	428	10		
1	B	279	Total	C	N	O	S	11	0
			2054	1255	361	428	10		


There are 2 discrepancies between the modelled and reference sequences:

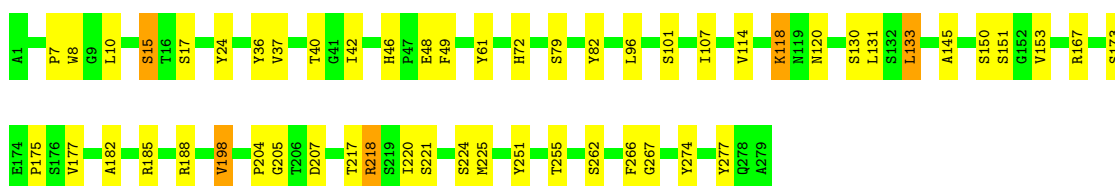
Chain	Residue	Modelled	Actual	Comment	Reference
A	207	ASP	SER	conflict	UNP P06873
B	207	ASP	SER	conflict	UNP P06873

3 Residue-property plots


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. 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. 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.

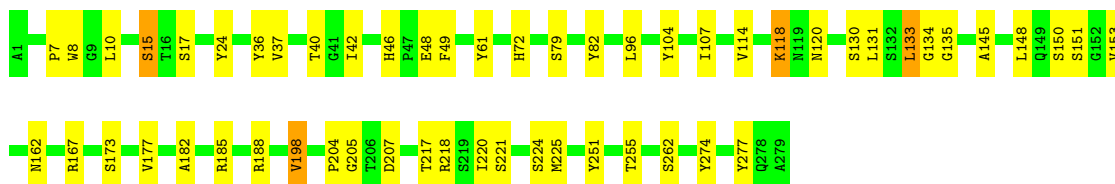
- Molecule 1: Proteinase K

Chain A: 



- Molecule 1: Proteinase K

Chain B: 



4 Data and refinement statistics i

EDS was not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	38.33Å 128.66Å 48.04Å 90.00° 90.52° 90.00°	Depositor
Resolution (Å)	2.50 – 2.50	Depositor
% Data completeness (in resolution range)	81.8 (2.50-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.44 (at 2.48Å)	Xtrriage
Refinement program	unknown	Depositor
R, R_{free}	0.217 , 0.262	Depositor
Wilson B-factor (Å ²)	43.5	Xtrriage
Anisotropy	0.048	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.087 for h,-k,-l	Xtrriage
Reported twinning fraction	0.716 for H, K, L 0.284 for -h,-k,l	Depositor
Outliers	0 of 13645 reflections	Xtrriage
Total number of atoms	4108	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.53% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.43	0/2149	0.81	2/2917 (0.1%)
1	B	0.43	0/2149	0.82	1/2917 (0.0%)
All	All	0.43	0/4298	0.82	3/5834 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	167	ARG	CB-CG-CD	5.42	123.78	111.30
1	B	120	ASN	CB-CA-C	5.21	119.47	111.02
1	A	120	ASN	CB-CA-C	5.17	119.39	111.02

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	2054	0	1953	39	0
1	B	2054	0	1953	38	0
All	All	4108	0	3906	70	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 9.

The worst 5 of 70 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:267:GLY:HA2	1:B:135:GLY:HA3	1.38	1.00
1:A:173:SER:HA	1:A:198:VAL:HG21	1.55	0.89
1:B:173:SER:HA	1:B:198:VAL:HG21	1.56	0.88
1:A:130:SER:C	1:A:131:LEU:HD12	2.09	0.78
1:B:130:SER:C	1:B:131:LEU:HD12	2.09	0.77

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 PDB entries followed by that with respect to all EM entries.

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	288/279 (103%)	277 (96%)	11 (4%)	0	100	100
1	B	288/279 (103%)	277 (96%)	11 (4%)	0	100	100
All	All	576/558 (103%)	554 (96%)	22 (4%)	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 PDB entries followed by that with respect to all EM entries.

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	224/213 (105%)	216 (96%)	8 (4%)	31	58
1	B	224/213 (105%)	217 (97%)	7 (3%)	35	62
All	All	448/426 (105%)	433 (97%)	15 (3%)	32	61

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	218	ARG
1	B	167	ARG
1	B	15	SER
1	B	198	VAL
1	B	133	LEU

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	119	ASN
1	B	119	ASN
1	B	168	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](#)

Mogul was not executed - this section is therefore empty.

5.5 Carbohydrates [i](#)

Mogul was not executed - this section is therefore empty.

5.6 Ligand geometry [i](#)

Mogul was not executed - this section is therefore empty.

5.7 Other polymers [i](#)

Mogul was not executed - this section is therefore empty.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.