



wwPDB EM Validation Summary Report ⓘ

Jun 22, 2026 – 12:15 PM EDT

PDB ID : 9OYY / pdb_00009oyy
Title : MicroED structure of proteinase K from microcrystals frozen by traditional cryoEM methods
Authors : Vlahakis, N.; Summers, J.A.; Rodriguez, J.A.; Dahlberg, P.; Wakatsuki, S.
Deposited on : 2025-06-04
Resolution : 2.20 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
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

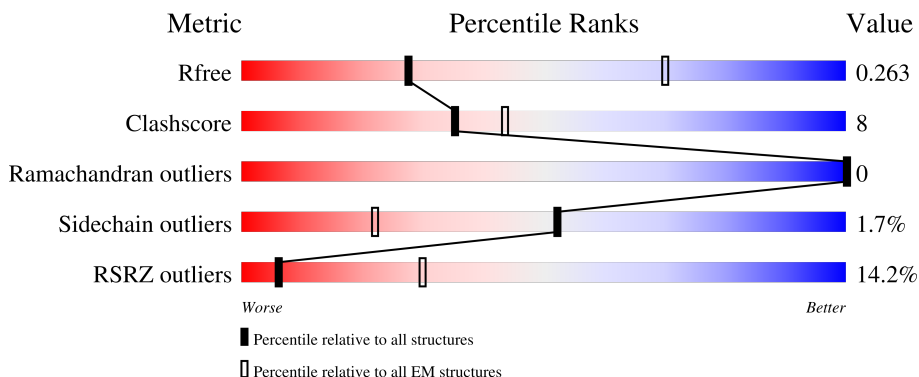
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON CRYSTALLOGRAPHY

The reported resolution of this entry is 2.20 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)
R_{free}	180332	208
Clashscore	229148	23984
Ramachandran outliers	224038	23583
Sidechain outliers	223484	23102
RSRZ outliers	180361	209

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	
1	B	279	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	CA	B	302	-	-	-	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4092 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	2037	1248	356	424	9	7	0
1	B	279	2033	1246	356	422	9	7	0

There are 2 discrepancies between the modelled and reference sequences:

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

- Molecule 2 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		AltConf
			Total	Ca	
2	A	2	2	2	0
2	B	2	2	2	0

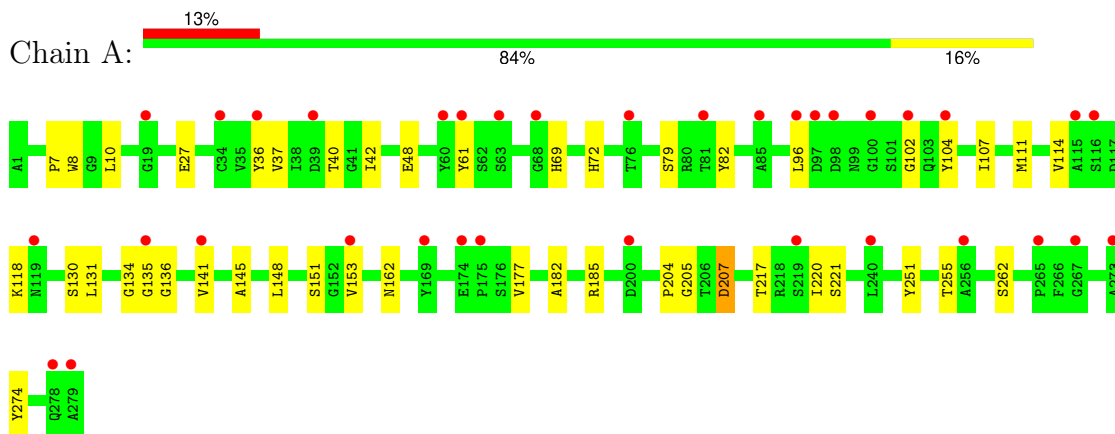
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		AltConf
			Total	O	
3	A	8	8	8	0
3	B	10	10	10	0

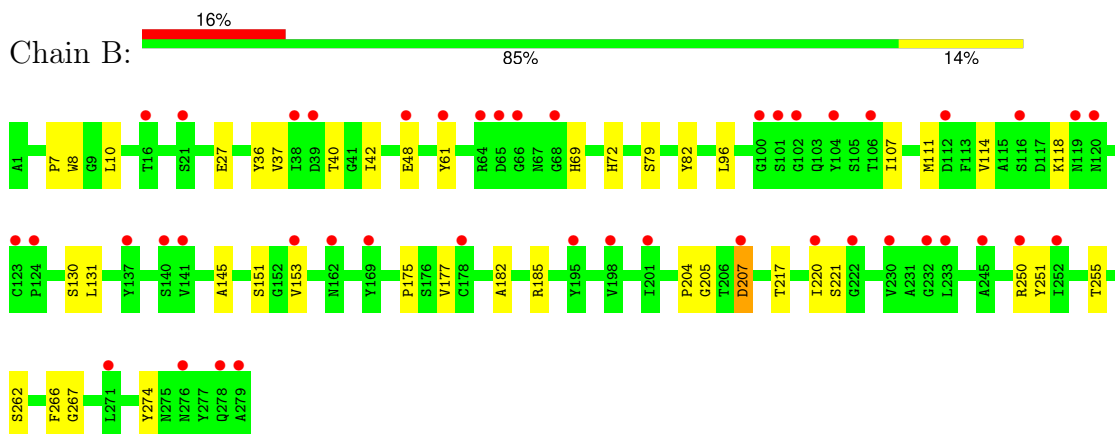
3 Residue-property plots [i](#)

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. 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



- Molecule 1: Proteinase K



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	38.30Å 129.10Å 47.70Å 90.00° 90.90° 90.00°	Depositor
Resolution (Å)	32.94 – 2.20 32.94 – 2.20	Depositor EDS
% Data completeness (in resolution range)	76.1 (32.94-2.20) 79.8 (32.94-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.40 (at 2.20Å)	Xtrriage
Refinement program	unknown	Depositor
R, R_{free}	0.210 , 0.237 0.245 , 0.263	Depositor DCC
R_{free} test set	962 reflections (4.09%)	wwPDB-VP
Wilson B-factor (Å ²)	26.4	Xtrriage
Anisotropy	0.083	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 22.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.084 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	4092	wwPDB-VP
Average B, all atoms (Å ²)	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.42% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	0.48	0/2111	0.81	2/2868 (0.1%)
1	B	0.50	0/2107	0.80	2/2862 (0.1%)
All	All	0.49	0/4218	0.81	4/5730 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	69	HIS	CA-CB-CG	5.39	119.19	113.80
1	B	207	ASP	CA-CB-CG	5.33	117.93	112.60
1	A	207	ASP	CA-CB-CG	5.31	117.91	112.60
1	B	69	HIS	CA-CB-CG	5.09	118.89	113.80

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	2037	0	1931	39	0
1	B	2033	0	1927	35	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	8	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	10	0	0	0	0
All	All	4092	0	3858	63	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 8.

The worst 5 of 63 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:135:GLY:HA3	1:B:267:GLY:HA2	1.37	1.05
1:B:72:HIS:CE1	1:B:220:ILE:HG12	2.31	0.66
1:A:72:HIS:CE1	1:A:220:ILE:HG12	2.31	0.65
1:A:118:LYS:HD3	1:A:153:VAL:CG2	2.29	0.62
1:A:104:TYR:CD1	1:A:141:VAL:HG21	2.35	0.61

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	284/279 (102%)	273 (96%)	11 (4%)	0	100	100
1	B	284/279 (102%)	274 (96%)	10 (4%)	0	100	100
All	All	568/558 (102%)	547 (96%)	21 (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	218/213 (102%)	215 (99%)	3 (1%)	59	75
1	B	217/213 (102%)	213 (98%)	4 (2%)	51	68
All	All	435/426 (102%)	428 (98%)	7 (2%)	52	71

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	27	GLU
1	B	61	TYR
1	B	262	SER
1	B	250	ARG
1	A	262	SER

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	162	ASN
1	B	162	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](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.