



## wwPDB EM Validation Summary Report i

Jul 6, 2025 – 12:09 AM JST

PDB ID : 9JAN / pdb\_00009jan  
EMDB ID : EMD-61294  
Title : Cryo-EM structure of MPXV protease in complex with compound A4  
Authors : Gao, Y.; Xie, X.; Lan, W.; Wang, W.; Yang, H.  
Deposited on : 2024-08-25  
Resolution : 2.93 Å (reported)

This is a wwPDB EM 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/EMValidationReportHelp>

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](#) i) were used in the production of this report:

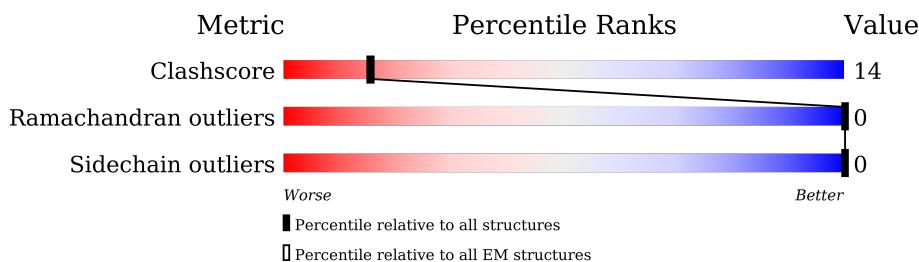
EMDB validation analysis : **FAILED**  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
MolProbity : 4-5-2 with Phenix2.0rc1  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.44

# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:  
**ELECTRON MICROSCOPY**

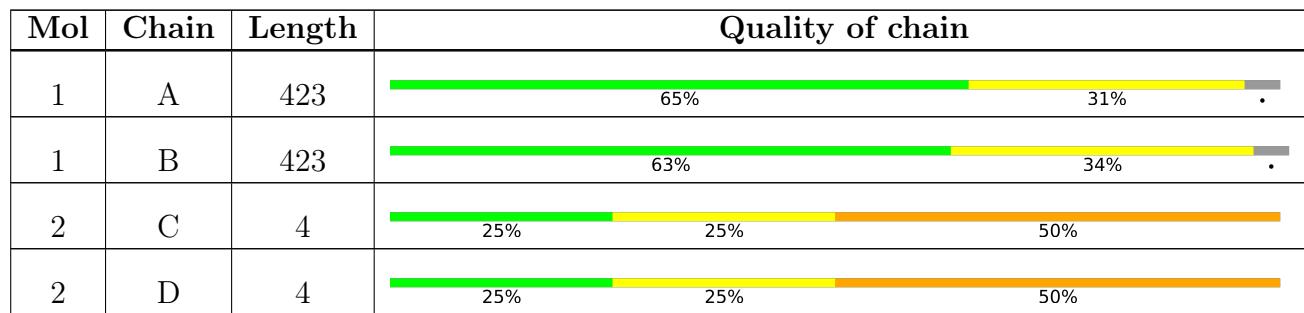
The reported resolution of this entry is 2.93 Å.

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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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  $\geq 3$ , 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  $\leq 5\%$



## 2 Entry composition [\(i\)](#)

There are 2 unique types of molecules in this entry. The entry contains 6550 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 Core protease I7.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	408	Total	C	N	O	S	0	0
			3237	2098	531	583	25		

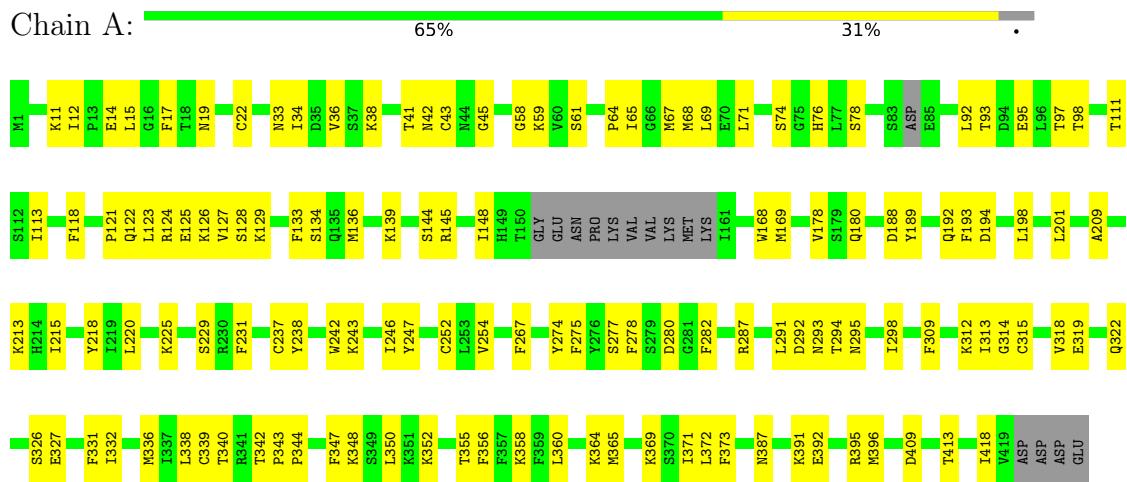
- Molecule 2 is a protein called A1ECL-DI8-ALA-AEM.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	4	Total	C	F	N	O	S	0
			38	25	2	5	5	1	

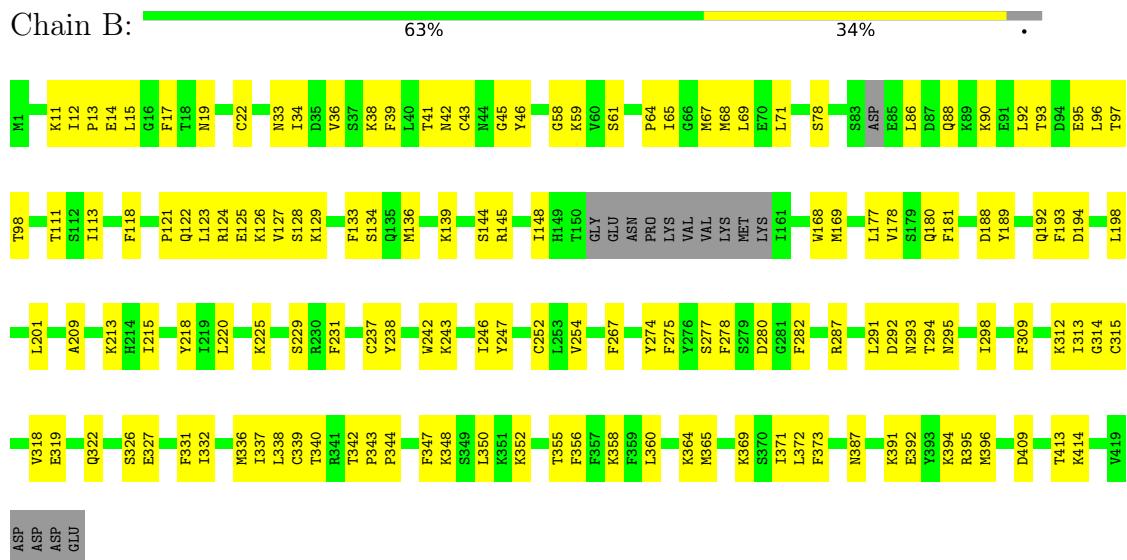
### 3 Residue-property plots [i](#)

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: Core protease I7



- Molecule 1: Core protease I7



- Molecule 2: A1ECL-DI8-ALA-AEM





- Molecule 2: A1ECL-DI8-ALA-AEM

Chain D:  25% 25% 50%



## 4 Experimental information i

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	717581	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	60	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	1800	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

## 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: DI8, A1ECL, AEM

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	0.32	0/3311	0.45	0/4460
1	B	0.33	0/3311	0.48	1/4460 (0.0%)
2	C	1.11	0/4	0.77	0/4
2	D	1.14	0/4	0.77	0/4
All	All	0.33	0/6630	0.46	1/8928 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	88	GLN	N-CA-C	9.49	121.71	111.36

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	3237	0	3160	90	0
1	B	3237	0	3160	99	0
2	C	38	0	17	5	0
2	D	38	0	17	5	0
All	All	6550	0	6354	187	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 14.

The worst 5 of 187 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:194:ASP:HB2	2:C:1:A1ECL:CE1	2.08	0.84
1:B:194:ASP:HB2	2:D:1:A1ECL:CE1	2.08	0.82
1:A:194:ASP:HB2	2:C:1:A1ECL:CZ	2.26	0.66
1:B:194:ASP:HB2	2:D:1:A1ECL:CZ	2.26	0.64
1:B:86:LEU:O	1:B:90:LYS:HG3	1.99	0.63

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	402/423 (95%)	367 (91%)	35 (9%)	0	100 100
1	B	402/423 (95%)	367 (91%)	35 (9%)	0	100 100
All	All	804/846 (95%)	734 (91%)	70 (9%)	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	351/393 (89%)	351 (100%)	0	100	100
1	B	351/393 (89%)	351 (100%)	0	100	100
All	All	702/786 (89%)	702 (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. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	81	ASN
1	B	180	GLN
1	B	321	ASN
1	B	192	GLN
1	A	293	ASN

### 5.3.3 RNA [\(i\)](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [\(i\)](#)

4 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	A1ECL	D	1	2	16,17,18	1.68	3 (18%)	24,24,26	2.99	7 (29%)
2	A1ECL	C	1	2	16,17,18	1.68	3 (18%)	24,24,26	2.99	7 (29%)
2	DI8	C	2	2	11,13,14	1.13	1 (9%)	13,17,19	0.59	0
2	DI8	D	2	2	11,13,14	1.14	1 (9%)	13,17,19	0.60	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns.

'-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	A1ECL	D	1	2	-	2/10/11/13	0/1/1/1
2	A1ECL	C	1	2	-	2/10/11/13	0/1/1/1
2	DI8	C	2	2	-	0/1/11/13	0/2/2/2
2	DI8	D	2	2	-	0/1/11/13	0/2/2/2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	A1ECL	S10-N	4.45	1.71	1.62
2	D	1	A1ECL	S10-N	4.43	1.71	1.62
2	D	1	A1ECL	C13-S10	2.88	1.82	1.75
2	C	1	A1ECL	C13-S10	2.87	1.82	1.75
2	D	1	A1ECL	CA-N	-2.40	1.45	1.48

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1	A1ECL	O12-S10-O11	-12.28	101.19	118.85
2	C	1	A1ECL	O12-S10-O11	-12.28	101.19	118.85
2	D	1	A1ECL	CD2-CE2-CZ	-2.78	120.01	123.52
2	C	1	A1ECL	CD2-CE2-CZ	-2.78	120.01	123.52
2	C	1	A1ECL	CD1-CE1-CZ	-2.77	120.02	123.52

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1	A1ECL	CA-N-S10-O12
2	C	1	A1ECL	CA-N-S10-C13
2	D	1	A1ECL	CA-N-S10-O12
2	D	1	A1ECL	CA-N-S10-C13

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1	A1ECL	2	0
2	C	1	A1ECL	2	0
2	C	2	DI8	1	0
2	D	2	DI8	1	0

## 5.5 Carbohydrates [\(i\)](#)

There are no oligosaccharides 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.