



# wwPDB EM Validation Summary Report i

Jul 6, 2025 – 12:12 AM JST

PDB ID : 9KR6 / pdb\_00009kr6  
EMDB ID : EMD-62520  
Title : Cryo-EM structure of MPXV core protease in complex with the substrate derivative I-G18  
Authors : Lan, W.; You, T.; Li, D.; Dong, X.; Wang, H.; Xu, J.; Wang, W.; Gao, Y.; Yang, H.  
Deposited on : 2024-11-27  
Resolution : 2.80 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

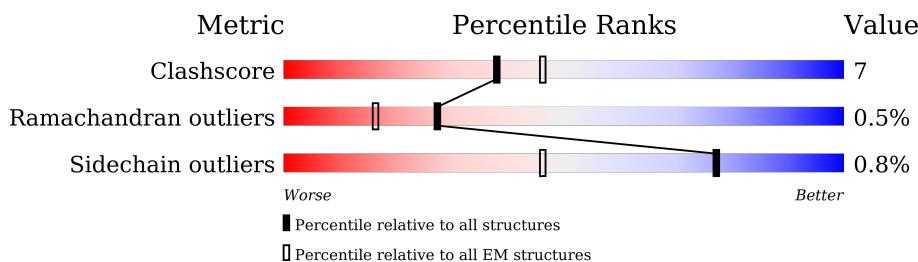
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *ELECTRON MICROSCOPY*

The reported resolution of this entry is 2.80 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 6760 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	409	Total	C	N	O	S	0	0
			3334	2153	543	613	25		

Mol	Chain	Residues	Atoms					AltConf	Trace
1	B	409	Total	C	N	O	S	0	0
			3334	2153	543	613	25		

- Molecule 2 is a protein called Core protein VP8.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	6	Total	C	N	O		0	0
			46	33	6	7			

Mol	Chain	Residues	Atoms					AltConf	Trace
2	D	6	Total	C	N	O		0	0
			46	33	6	7			

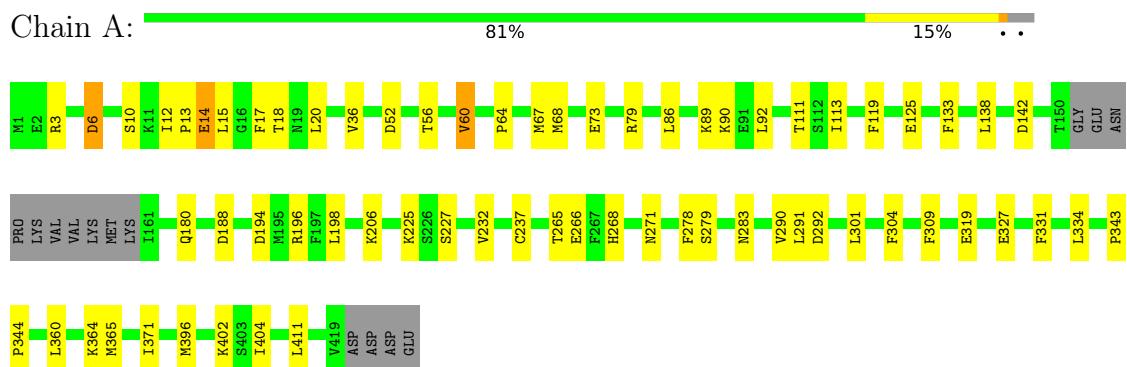
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	9	ETA	GLY	engineered mutation	UNP M1L511
D	9	ETA	GLY	engineered mutation	UNP M1L511

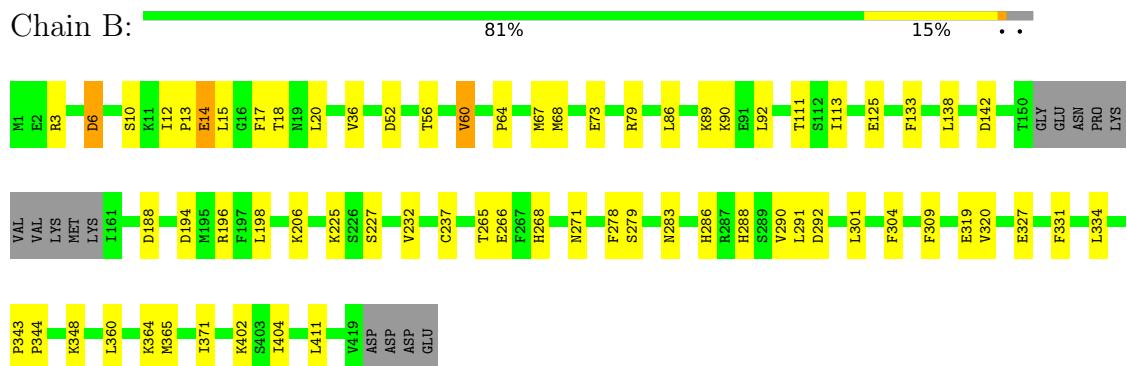
### 3 Residue-property plots [i](#)

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: Core protein VP8



- Molecule 2: Core protein VP8





## 4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	760440	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 ( $e^-/\text{\AA}^2$ )	60.00, 60	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	1800	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k), GATAN K3 BIO-QUANTUM (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: ETA

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	0/3410	0.51	0/4590
1	B	0.43	0/3410	0.51	0/4590
2	C	0.53	0/43	0.44	0/57
2	D	0.53	0/43	0.44	0/57
All	All	0.43	0/6906	0.51	0/9294

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
2	C	0	1
2	D	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	8	ALA	Mainchain
2	D	8	ALA	Mainchain

## 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	3334	0	3321	51	0
1	B	3334	0	3321	48	0
2	C	46	0	40	1	0
2	D	46	0	40	1	0
All	All	6760	0	6722	97	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 7.

The worst 5 of 97 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:14:GLU:HA	1:A:18:THR:HB	1.43	1.00
1:B:14:GLU:HA	1:B:18:THR:HB	1.43	0.97
1:B:14:GLU:HA	1:B:18:THR:CB	2.01	0.91
1:A:14:GLU:HA	1:A:18:THR:CB	2.01	0.88
1:B:13:PRO:O	1:B:15:LEU:N	2.19	0.76

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	405/423 (96%)	373 (92%)	30 (7%)	2 (0%)	25 56
1	B	405/423 (96%)	373 (92%)	30 (7%)	2 (0%)	25 56

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	C	4/10 (40%)	4 (100%)	0	0	100	100
2	D	4/10 (40%)	3 (75%)	1 (25%)	0	100	100
All	All	818/866 (94%)	753 (92%)	61 (8%)	4 (0%)	27	56

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	14	GLU
1	B	14	GLU
1	A	60	VAL
1	B	60	VAL

### 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	380/393 (97%)	377 (99%)	3 (1%)	79	93
1	B	380/393 (97%)	377 (99%)	3 (1%)	79	93
2	C	4/8 (50%)	4 (100%)	0	100	100
2	D	4/8 (50%)	4 (100%)	0	100	100
All	All	768/802 (96%)	762 (99%)	6 (1%)	77	93

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	6	ASP
1	B	232	VAL
1	B	411	LEU
1	A	232	VAL
1	A	6	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	322	GLN
1	B	387	ASN
1	A	387	ASN
1	B	44	ASN
1	B	88	GLN

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

There are no RNA molecules in this entry.

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

2 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	ETA	D	9	2	3,3,3	0.48	0	2,2,2	0.70	0
2	ETA	C	9	2	3,3,3	0.49	0	2,2,2	0.71	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	ETA	D	9	2	-	1/1/1/1	-
2	ETA	C	9	2	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	9	ETA	O-C-CA-N
2	D	9	ETA	O-C-CA-N

There are no ring outliers.

No monomer is involved in short contacts.

## 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.