wwPDB EM Validation Summary Report

May 14, 2024 – 03:00 am BST

PDB ID	:	8P1N
EMDB ID	:	EMD-17355
Title	:	Structure of hantaan orthohantavirus (HTNV) polymerase bound to 5'vRNA
Authors	:	Keown, J.R.; Carrique, L.; Grimes, J.M.
Deposited on	:	2023-05-12
Resolution	:	2.79 Å(reported)
Based on initial model	:	.

This is a wwPDB EM Validation Summary 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/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:

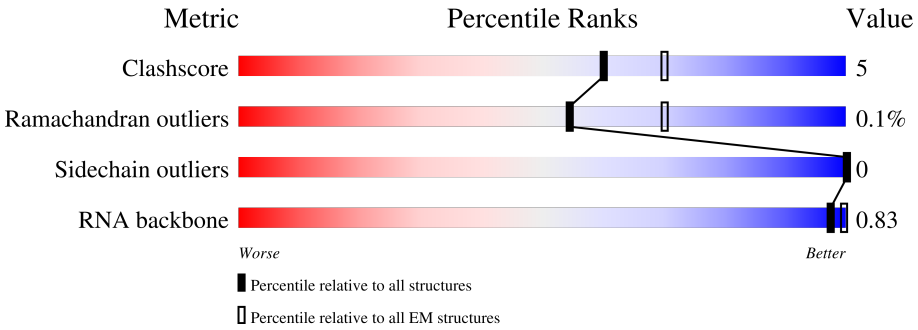
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.79 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	2196	52% 7% 41%
2	B	17	18% 12% 71%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 20875 atoms, of which 10405 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 RNA-directed RNA polymerase L.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
1	A	1287	20765	6669	10405	1738	1897	56	0	0

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-44	MET	-	initiating methionine	UNP P23456
A	-43	TRP	-	expression tag	UNP P23456
A	-42	SER	-	expression tag	UNP P23456
A	-41	HIS	-	expression tag	UNP P23456
A	-40	PRO	-	expression tag	UNP P23456
A	-39	GLN	-	expression tag	UNP P23456
A	-38	PHE	-	expression tag	UNP P23456
A	-37	GLU	-	expression tag	UNP P23456
A	-36	LYS	-	expression tag	UNP P23456
A	-35	GLY	-	expression tag	UNP P23456
A	-34	GLY	-	expression tag	UNP P23456
A	-33	GLY	-	expression tag	UNP P23456
A	-32	SER	-	expression tag	UNP P23456
A	-31	GLY	-	expression tag	UNP P23456
A	-30	GLY	-	expression tag	UNP P23456
A	-29	GLY	-	expression tag	UNP P23456
A	-28	SER	-	expression tag	UNP P23456
A	-27	GLY	-	expression tag	UNP P23456
A	-26	GLY	-	expression tag	UNP P23456
A	-25	SER	-	expression tag	UNP P23456
A	-24	SER	-	expression tag	UNP P23456
A	-23	ALA	-	expression tag	UNP P23456
A	-22	TRP	-	expression tag	UNP P23456
A	-21	SER	-	expression tag	UNP P23456
A	-20	HIS	-	expression tag	UNP P23456
A	-19	PRO	-	expression tag	UNP P23456
A	-18	GLN	-	expression tag	UNP P23456
A	-17	PHE	-	expression tag	UNP P23456

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-16	GLU	-	expression tag	UNP P23456
A	-15	LYS	-	expression tag	UNP P23456
A	-14	HIS	-	expression tag	UNP P23456
A	-13	HIS	-	expression tag	UNP P23456
A	-12	HIS	-	expression tag	UNP P23456
A	-11	HIS	-	expression tag	UNP P23456
A	-10	HIS	-	expression tag	UNP P23456
A	-9	HIS	-	expression tag	UNP P23456
A	-8	HIS	-	expression tag	UNP P23456
A	-7	HIS	-	expression tag	UNP P23456
A	-6	GLU	-	expression tag	UNP P23456
A	-5	ASN	-	expression tag	UNP P23456
A	-4	LEU	-	expression tag	UNP P23456
A	-3	TYR	-	expression tag	UNP P23456
A	-2	PHE	-	expression tag	UNP P23456
A	-1	GLN	-	expression tag	UNP P23456
A	0	GLY	-	expression tag	UNP P23456
A	97	ALA	ASP	engineered mutation	UNP P23456

- Molecule 2 is a RNA chain called RNA (5'-R(P*AP*GP*UP*AP*G)-3').

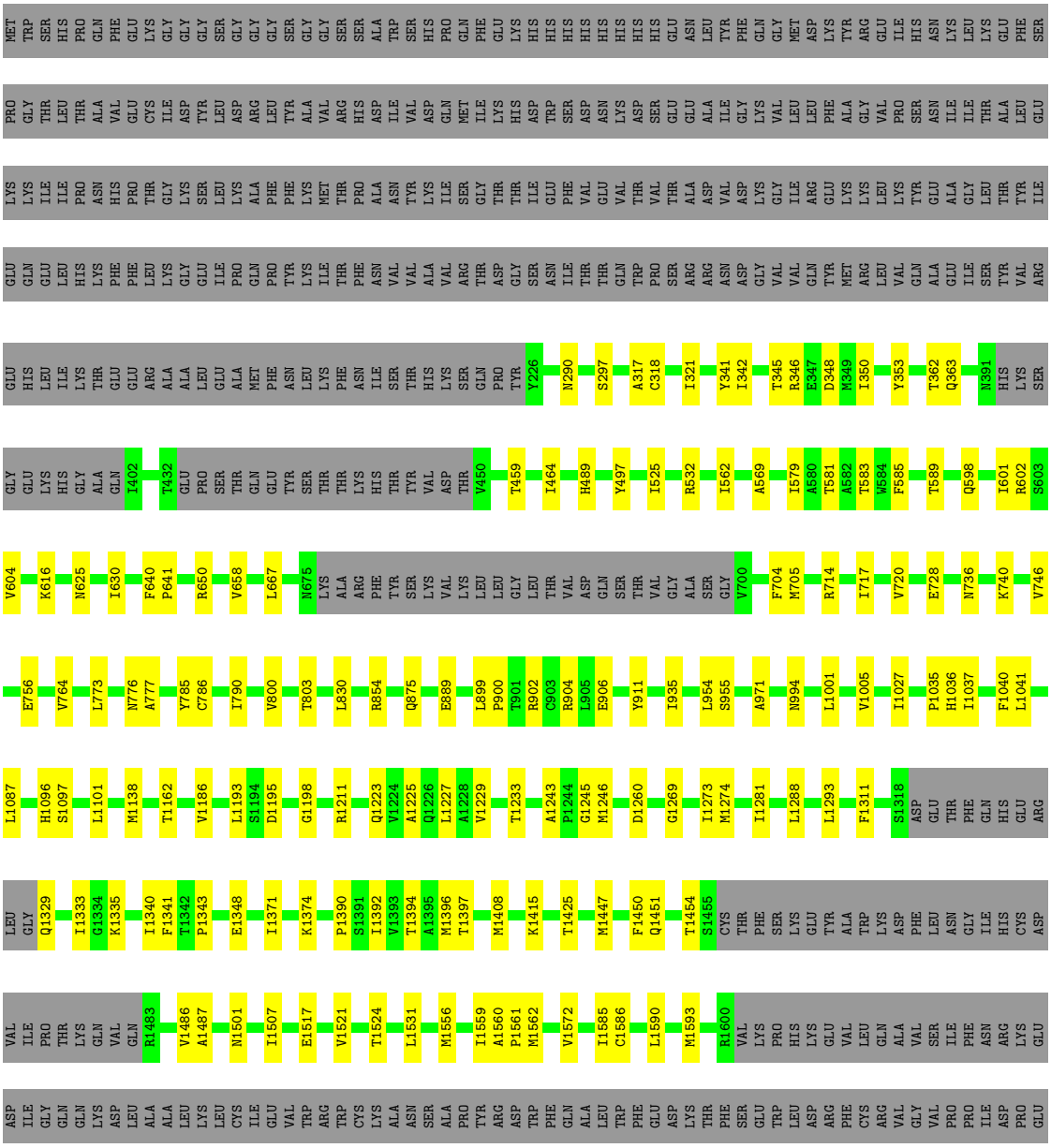
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	B	5	110	49	22	34	5	0	0

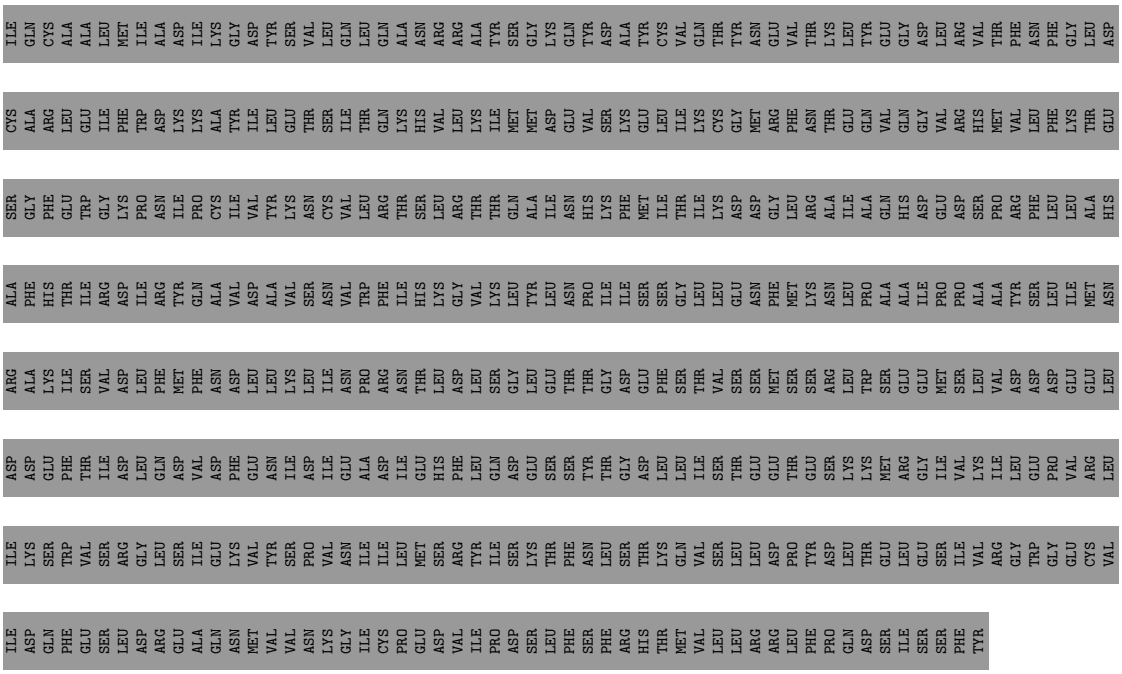
3 Residue-property plots (i)

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. 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: RNA-directed RNA polymerase L

Chain A:





• Molecule 2: RNA (5'-R(P*AP*GP*UP*AP*G)-3')

Chain B: 18% 12% 71%

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	67000	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 ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1400	Depositor
Maximum defocus (nm)	2600	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor

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	0.25	0/10589	0.46	0/14305
2	B	0.22	0/123	0.62	0/190
All	All	0.25	0/10712	0.46	0/14495

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	10360	10405	10405	97	0
2	B	110	0	55	1	0
All	All	10470	10405	10460	98	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 5.

The worst 5 of 98 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:341:TYR:HH	1:A:353:TYR:HH	1.15	0.85
1:A:650:ARG:NH2	1:A:889:GLU:OE1	2.13	0.80

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:583:THR:HG21	1:A:1186:VAL:CG2	2.19	0.72
1:A:625:ASN:OD1	1:A:650:ARG:NH2	2.24	0.68
1:A:630:ILE:CD1	1:A:720:VAL:HG21	2.24	0.67

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	1275/2196 (58%)	1224 (96%)	50 (4%)	1 (0%)	51 81

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1097	SER

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	1140/1955 (58%)	1140 (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. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	474	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	4/17 (23%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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 monosaccharides 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.

6 Map visualisation

This section contains visualisations of the EMDB entry EMD-17355. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

6.5 Orthogonal surface views

This section was not generated.

6.6 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution

This section was not generated.

7.2 Volume estimate versus contour level

This section was not generated.

7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.