



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jun 23, 2018 – 10:20 PM EDT

PDB ID : 6F5O
EMDB ID: : EMD-4190
Title : A mechanism for the activation of the influenza virus transcriptase
Authors : Serna Martin, I.; Grimes, J.M.
Deposited on : 2017-12-02
Resolution : 9.80 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model 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.

MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031172

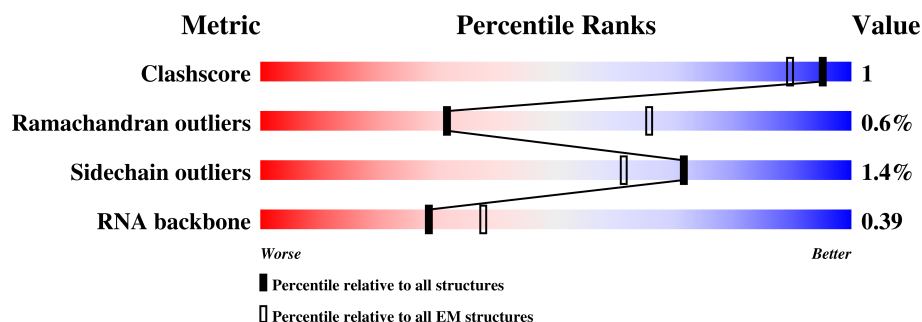
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 9.80 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	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531
RNA backbone	3747	458

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	727	94% . .
2	B	752	94% . .
3	C	770	88% 7% 5%
4	R	14	64% 36%
5	V	14	79% 21%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 17940 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 Polymerase acidic protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	716	Total	C	N	O	S	0	0
			5747	3651	962	1094	40		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP Q5V8Z9

- Molecule 2 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	733	Total	C	N	O	S	0	0
			5761	3636	999	1074	52		

- Molecule 3 is a protein called Polymerase basic protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	730	Total	C	N	O	S	0	0
			5843	3715	1023	1065	40		

- Molecule 4 is a RNA chain called 3' promoter vRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	R	14	Total	C	N	O	P	0	0
			282	128	42	99	13		

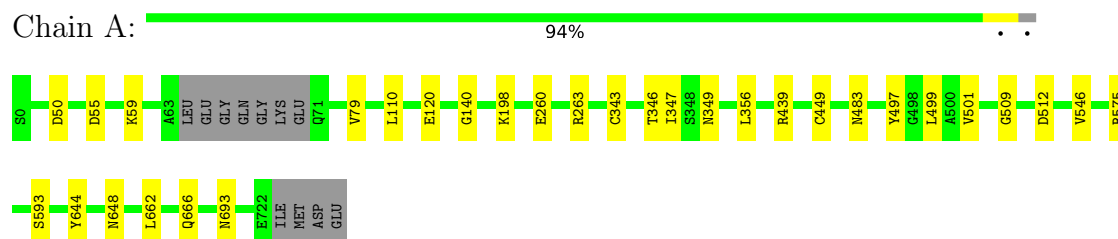
- Molecule 5 is a RNA chain called 5' promoter vRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	V	14	Total	C	N	O	P	0	0
			307	137	62	94	14		

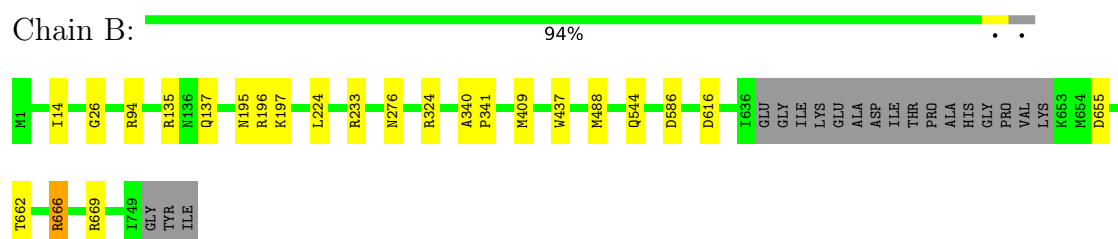
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

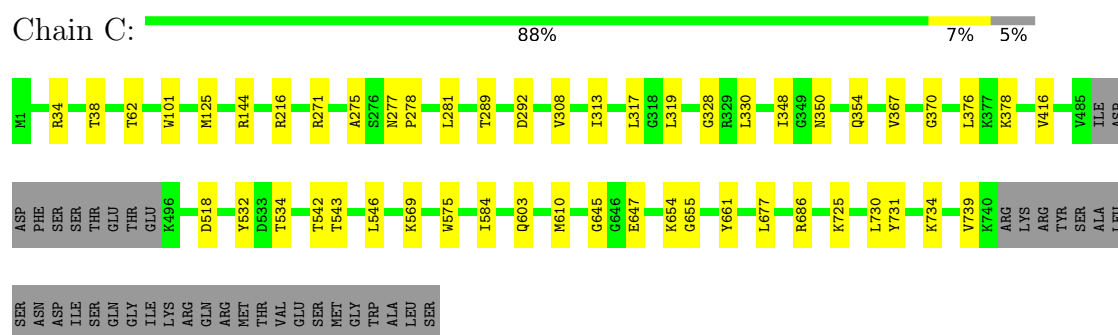
- Molecule 1: Polymerase acidic protein



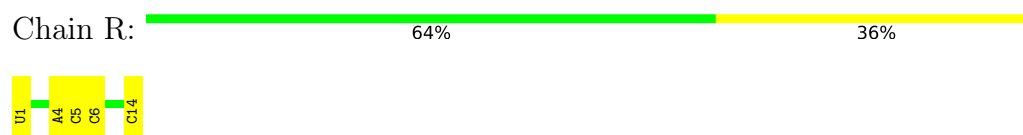
- Molecule 2: RNA-directed RNA polymerase catalytic subunit




- Molecule 3: Polymerase basic protein 2



- Molecule 4: 3' promoter vRNA



- Molecule 5: 5' promoter vRNA

Chain V:  79% 21%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	9159	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	2.09	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.23	0/5862	0.38	0/7904
2	B	0.23	0/5873	0.39	0/7915
3	C	0.23	0/5943	0.40	0/7987
4	R	0.26	0/312	0.81	0/482
5	V	0.62	1/345 (0.3%)	0.70	0/535
All	All	0.24	1/18335 (0.0%)	0.41	0/24823

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	V	1	A	OP3-P	-10.62	1.48	1.61

There are no bond angle outliers.

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	5747	0	5718	13	0
2	B	5761	0	5783	14	0
3	C	5843	0	6015	21	0
4	R	282	0	146	0	0
5	V	307	0	153	0	0
All	All	17940	0	17815	39	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 1.

The worst 5 of 39 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)
2:B:195:ASN:O	2:B:197:LYS:N	2.32	0.62
1:A:648:ASN:OD1	1:A:693:ASN:ND2	2.40	0.54
1:A:346:THR:O	1:A:349:ASN:ND2	2.40	0.54
1:A:509:GLY:N	1:A:512:ASP:OD2	2.42	0.52
1:A:666:GLN:HB2	2:B:14:ILE:HD12	1.93	0.51

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	710/727 (98%)	666 (94%)	41 (6%)	3 (0%)	36	77
2	B	729/752 (97%)	689 (94%)	39 (5%)	1 (0%)	53	88
3	C	726/770 (94%)	676 (93%)	42 (6%)	8 (1%)	16	58
All	All	2165/2249 (96%)	2031 (94%)	122 (6%)	12 (1%)	31	70

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	196	ARG
3	C	292	ASP
3	C	569	LYS
3	C	730	LEU
1	A	50	ASP

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	637/646 (99%)	629 (99%)	8 (1%)	71	86
2	B	631/645 (98%)	624 (99%)	7 (1%)	76	88
3	C	638/674 (95%)	627 (98%)	11 (2%)	63	83
All	All	1906/1965 (97%)	1880 (99%)	26 (1%)	71	85

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	437	TRP
2	B	666	ARG
3	C	610	MET
2	B	488	MET
2	B	616	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	666	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	R	13/14 (92%)	4 (30%)	1 (7%)
5	V	13/14 (92%)	2 (15%)	0
All	All	26/28 (92%)	6 (23%)	1 (3%)

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	R	4	A
4	R	5	C
4	R	6	C

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	R	14	C
5	V	7	A

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	R	1	U

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 carbohydrates 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	164:GLU	C	165:GLU	N	3.00