



wwPDB/EMDatabank EM Map/Model Validation Summary Report

Sep 19, 2019 – 12:00 PM EDT


PDB ID : 6QWL
EMDB ID: : EMD-4660
Title : Influenza B virus (B/Panama/45) polymerase Heterotrimer in complex with 3'5' cRNA promoter
Authors : Keown, J.R.; Carrique, L.; Fan, H.; Fodor, E.; Grimes, J.M.
Deposited on : 2019-03-05
Resolution : 4.10 Å (reported)
Based on PDB ID : 5EPI

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) : 2.4

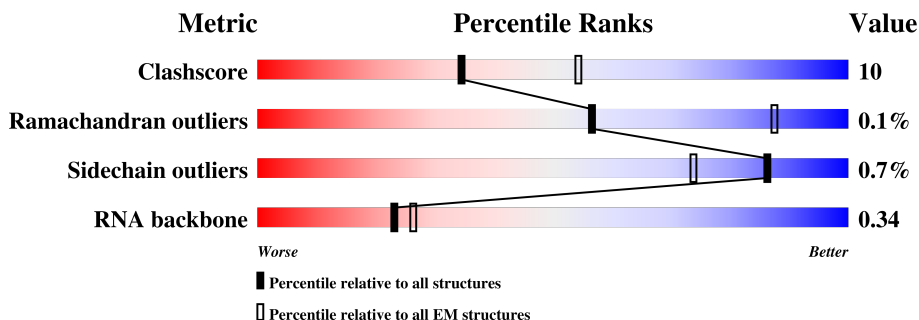
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.10 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	E	726	55% (green), 15% (yellow), 30% (grey)
2	K	752	53% (green), 19% (yellow), 28% (grey)
3	Q	778	18% (green), 7% (yellow), 75% (grey)
4	T	15	13% (green), 60% (yellow), 7% (orange), 20% (grey)
5	W	14	29% (green), 57% (yellow), 14% (orange)

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 10408 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
			Total	C	N	O	S		
1	E	511	4084	2593	696	765	30	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	K	539	4206	2656	717	799	34	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	Q	194	1565	1002	276	277	10	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	771	ALA	-	expression tag	UNP O36431
Q	772	ARG	-	expression tag	UNP O36431
Q	773	GLU	-	expression tag	UNP O36431
Q	774	ASN	-	expression tag	UNP O36431
Q	775	LEU	-	expression tag	UNP O36431
Q	776	TYR	-	expression tag	UNP O36431
Q	777	PHE	-	expression tag	UNP O36431
Q	778	GLN	-	expression tag	UNP O36431

- Molecule 4 is a RNA chain called 3' cRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
4	T	12	249	111	36	90	12	0	0

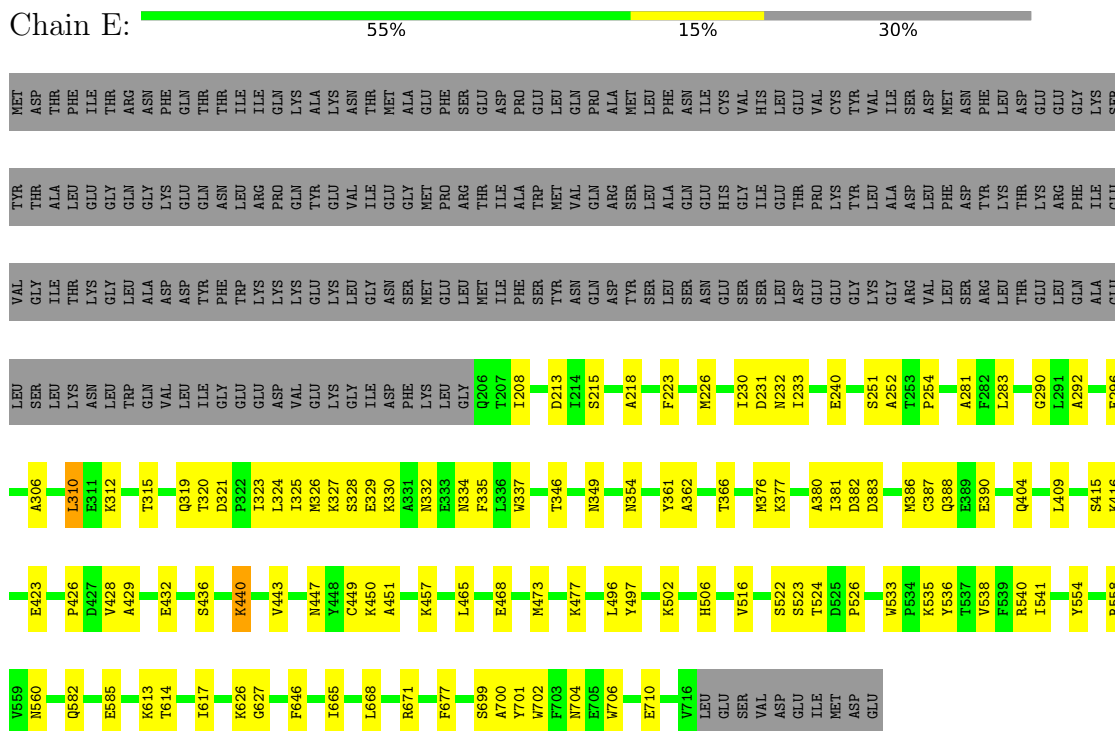
- Molecule 5 is a RNA chain called 5' cRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
5	W	14	304	136	62	92	14	0	0

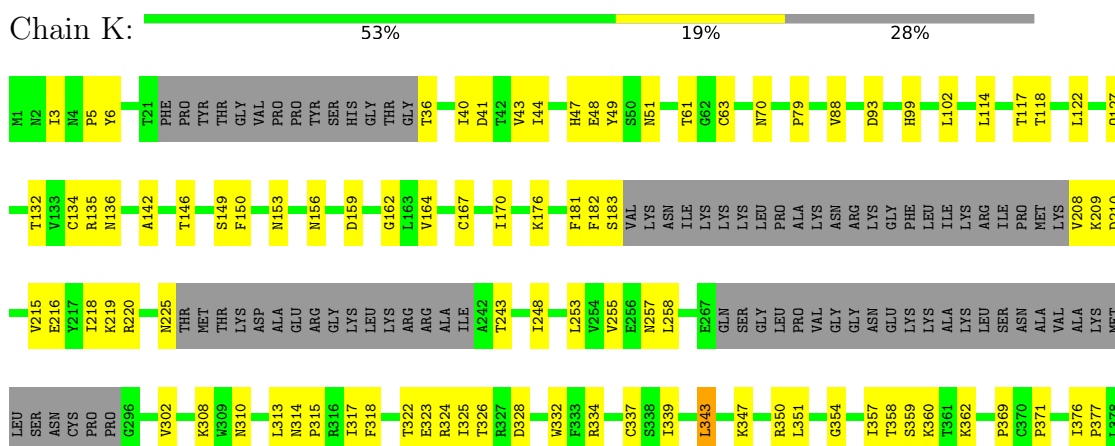
3 Residue-property plots [i](#)

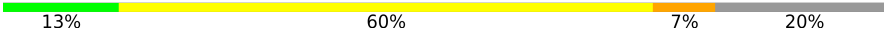
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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



Chain T:  13% 60% 7% 20%



• Molecule 5: 5' cRNA

Chain W:  29% 57% 14%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	1012085	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	1.1	Depositor
Minimum defocus (nm)	1300	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	130000	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	E	0.49	0/4170	0.59	0/5624
2	K	0.39	0/4282	0.59	0/5785
3	Q	0.30	0/1600	0.54	0/2159
4	T	0.67	0/275	1.43	8/425 (1.9%)
5	W	0.70	0/341	1.33	2/530 (0.4%)
All	All	0.44	0/10668	0.66	10/14523 (0.1%)

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
4	T	3	C	C5-C6-N1	7.36	124.68	121.00
4	T	3	C	C6-N1-C2	-6.07	117.87	120.30
4	T	2	G	C4-N9-C1'	6.06	134.38	126.50
4	T	2	G	C6-C5-N7	-5.85	126.89	130.40
4	T	2	G	C8-N9-C1'	-5.70	119.59	127.00

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	E	4084	0	4086	82	0
2	K	4206	0	4176	106	0
3	Q	1565	0	1598	37	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	T	249	0	127	0	0
5	W	304	0	155	3	0
All	All	10408	0	10142	197	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 10.

The worst 5 of 197 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:E:386:MET:HB3	2:K:357:ILE:HD12	1.71	0.72
2:K:517:ASN:HD21	2:K:520:ALA:H	1.35	0.71
2:K:99:HIS:HB3	2:K:102:LEU:HB2	1.73	0.70
1:E:526:PRO:HG3	1:E:538:VAL:HG11	1.78	0.65
1:E:449:CYS:SG	1:E:450:LYS:N	2.70	0.64

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	E	509/726 (70%)	445 (87%)	64 (13%)	0	100	100
2	K	525/752 (70%)	464 (88%)	60 (11%)	1 (0%)	49	84
3	Q	192/778 (25%)	183 (95%)	9 (5%)	0	100	100
All	All	1226/2256 (54%)	1092 (89%)	133 (11%)	1 (0%)	56	87

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	K	443	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	E	452/645 (70%)	449 (99%)	3 (1%)	85	92
2	K	461/644 (72%)	458 (99%)	3 (1%)	85	92
3	Q	166/682 (24%)	164 (99%)	2 (1%)	74	87
All	All	1079/1971 (55%)	1071 (99%)	8 (1%)	86	92

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	K	164	VAL
3	Q	156	MET
2	K	604	HIS
1	E	646	PHE
2	K	343	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	704	ASN
2	K	51	ASN
2	K	261	ASN
1	E	666	GLN
2	K	257	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	T	11/15 (73%)	9 (81%)	0
5	W	13/14 (92%)	6 (46%)	0
All	All	24/29 (82%)	15 (62%)	0

5 of 15 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	T	3	C
4	T	4	C
4	T	5	U
4	T	6	U
4	T	7	G

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 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](#)

There are no chain breaks in this entry.