



# wwPDB/EMDatabank EM Map/Model Validation Summary Report

Sep 19, 2019 – 10:27 AM EDT


PDB ID : 6QXE  
EMDB ID: : EMD-4666  
Title : Influenza A virus (A/NT/60/1968) polymerase dimer of heterotrimer in complex with 3'5' cRNA promoter and Nb8205  
Authors : Carrique, L.; Keown, J.R.; Fan, H.; Fodor, E.; Grimes, J.M.  
Deposited on : 2019-03-07  
Resolution : 4.15 Å (reported)  
Based on PDB ID : 6QPG, 6QX8

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](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  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

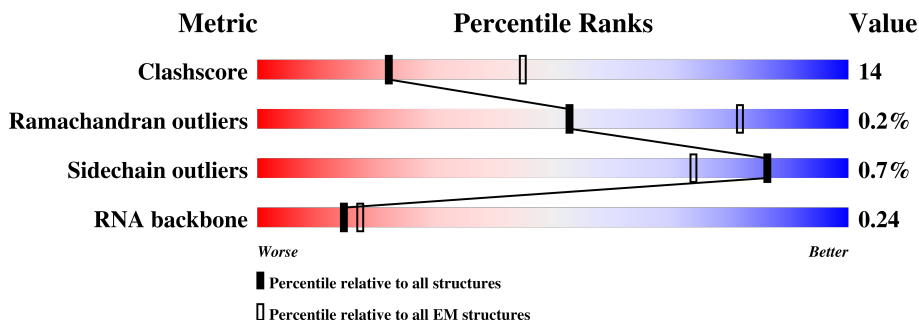
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.15 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.



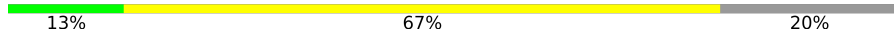
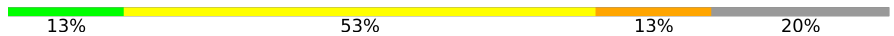
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531
RNA backbone	3747	458

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  $\geq 3$ , 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	M	134	
1	O	134	
2	A	716	
2	E	716	
3	B	757	
3	F	757	
4	C	759	
4	G	759	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	D	15	
5	H	15	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 40460 atoms, of which 19050 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 Nb8205.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	M	121	933	581	164	182	6	0	0
1	O	121	933	581	164	182	6	0	0

- Molecule 2 is a protein called Polymerase acidic protein.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
2	A	514	8141	2605	4035	693	779	29	0	0
2	E	514	8141	2605	4035	693	779	29	0	0

- Molecule 3 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
3	B	611	9631	3051	4791	838	916	35	0	0
3	F	611	9631	3051	4791	838	916	35	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	577	LYS	GLU	conflict	UNP P03432
F	577	LYS	GLU	conflict	UNP P03432

- Molecule 4 is a protein called Polymerase basic protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
4	C	67	1129	362	567	99	98	3	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
4	G	67	1129	362	567	99	98	3	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			P
5	D	12	396	118	132	56	78	12	0	0
5	H	12	396	118	132	56	78	12	0	0











GLY  
LYS  
GLU  
ASP  
ARG  
ARG  
TYR  
GLY  
PRO  
ALA  
LEU  
SER  
ILE  
ASN  
GLU  
LEU  
SER  
ASN  
LEU  
ALA  
LYS  
GLY  
GLU  
LYS  
ALA  
ASN  
VAL  
LEU  
ILE  
GLY  
GLN  
GLY  
ASP  
VAL  
VAL  
LEU  
VAL  
MET  
LYS  
ARG  
LYS  
ARG  
ASP  
SER  
SER  
ILE  
LEU  
THR  
ASP  
SER  
GLN  
THR  
ALA  
THR  
LYS  
ARG  
ILE  
ARG  
MET  
ALA

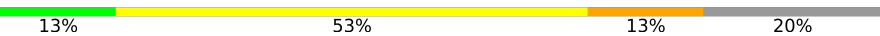
ILE  
ASN

- Molecule 5: RNA (5'-R(P\*AP\*GP\*CP\*AP\*AP\*AP\*AP\*GP\*CP\*AP\*GP\*G)-3')

Chain D:  13% 67% 20%

P A1 G2 C3 A4 A5 A6 A7 G8 C9 A10 G11 G12 C C

- Molecule 5: RNA (5'-R(P\*AP\*GP\*CP\*AP\*AP\*AP\*AP\*GP\*CP\*AP\*GP\*G)-3')

Chain H:  13% 53% 13% 20%

P A1 G2 C3 A4 A5 A6 A7 G8 C9 A10 G11 G12 C C

## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	27861	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	1.25	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	700	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	M	0.25	0/954	0.47	0/1293
1	O	0.26	0/954	0.48	0/1293
2	A	0.46	0/4196	0.59	1/5670 (0.0%)
2	E	0.46	0/4196	0.59	2/5670 (0.0%)
3	B	0.39	0/4938	0.61	1/6676 (0.0%)
3	F	0.39	0/4938	0.59	0/6676
4	C	0.37	0/577	0.58	0/781
4	G	0.37	0/577	0.58	0/781
5	D	0.71	0/297	1.07	0/462
5	H	0.73	0/297	1.07	0/462
All	All	0.42	0/21924	0.60	4/29764 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	34	THR	C-N-CA	8.82	140.83	122.30
2	E	213	ARG	NE-CZ-NH1	5.25	122.92	120.30
2	A	213	ARG	NE-CZ-NH1	5.02	122.81	120.30
2	E	716	ARG	NE-CZ-NH1	5.01	122.80	120.30

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	M	933	0	893	81	0
1	O	933	0	893	94	0
2	A	4106	4035	4033	128	0
2	E	4106	4035	4033	131	0
3	B	4840	4791	4791	146	0
3	F	4840	4791	4791	159	0
4	C	562	567	566	12	0
4	G	562	567	566	22	0
5	D	264	132	133	5	0
5	H	264	132	133	4	0
All	All	21410	19050	20832	583	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 583 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:O:43:LYS:CD	2:E:554:ILE:HG13	1.42	1.44
1:O:105:TRP:CB	3:B:587:ALA:CB	1.96	1.41
1:O:105:TRP:HB3	3:B:587:ALA:CB	1.53	1.38
1:O:105:TRP:HH2	3:B:563:ARG:NH1	1.14	1.38
1:O:43:LYS:HD3	2:E:554:ILE:CG1	1.55	1.37

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	M	119/134 (89%)	115 (97%)	4 (3%)	0	100 100
1	O	119/134 (89%)	116 (98%)	3 (2%)	0	100 100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	A	512/716 (72%)	475 (93%)	36 (7%)	1 (0%)	49	83
2	E	512/716 (72%)	475 (93%)	36 (7%)	1 (0%)	49	83
3	B	605/757 (80%)	551 (91%)	53 (9%)	1 (0%)	49	83
3	F	605/757 (80%)	551 (91%)	53 (9%)	1 (0%)	49	83
4	C	63/759 (8%)	58 (92%)	5 (8%)	0	100	100
4	G	63/759 (8%)	57 (90%)	6 (10%)	0	100	100
All	All	2598/4732 (55%)	2398 (92%)	196 (8%)	4 (0%)	53	83

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	571	SER
3	B	510	PRO
2	E	571	SER
3	F	510	PRO

### 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	M	99/111 (89%)	98 (99%)	1 (1%)	78	88
1	O	99/111 (89%)	97 (98%)	2 (2%)	58	79
2	A	451/644 (70%)	447 (99%)	4 (1%)	81	90
2	E	451/644 (70%)	447 (99%)	4 (1%)	81	90
3	B	531/669 (79%)	527 (99%)	4 (1%)	83	91
3	F	531/669 (79%)	530 (100%)	1 (0%)	94	96
4	C	62/670 (9%)	62 (100%)	0	100	100
4	G	62/670 (9%)	62 (100%)	0	100	100
All	All	2286/4188 (55%)	2270 (99%)	16 (1%)	86	92

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	28	PRO
3	B	54	LYS
2	E	493	GLU
2	A	716	ARG
2	E	551	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	B	621	GLN
3	F	621	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
5	D	11/15 (73%)	6 (54%)	1 (9%)
5	H	11/15 (73%)	6 (54%)	1 (9%)
All	All	22/30 (73%)	12 (54%)	2 (9%)

5 of 12 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
5	D	5	A
5	D	6	A
5	D	7	A
5	D	8	G
5	D	10	A

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
5	D	4	A
5	H	4	A

## 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.