



# wwPDB EM Model Validation Summary Report ⓘ

Apr 2, 2020 – 10:46 AM EDT

PDB ID : 6V5B  
EMDB ID : EMD-21051  
Title : Human Drosha and DGCR8 in complex with Primary MicroRNA (MP/RNA complex) - Active state  
Authors : Partin, A.; Zhang, K.; Jeong, B.; Herrell, E.; Li, S.; Chiu, W.; Nam, Y.  
Deposited on : 2019-12-04  
Resolution : 3.70 Å(reported)

This is a wwPDB EM 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.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
buster-report : 1.1.7 (2018)  
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) : 2.10.1

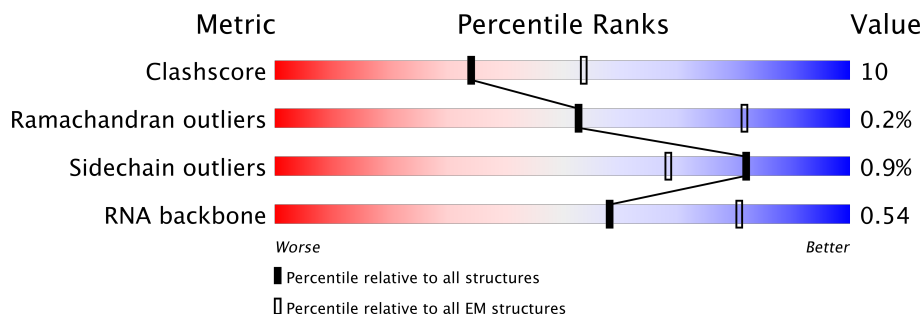
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.70 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 and their fit to the map. 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 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	1016	
2	B	532	
2	C	532	
3	D	105	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12563 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 Ribonuclease 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	908	Total	C	N	O	S	0	0
			7419	4726	1309	1334	50		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	350	GLY	-	expression tag	UNP Q9NRR4
A	351	SER	-	expression tag	UNP Q9NRR4
A	352	GLY	-	expression tag	UNP Q9NRR4

- Molecule 2 is a protein called Microprocessor complex subunit DGCR8.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	217	Total	C	N	O	S	0	0
			1750	1105	312	323	10		
2	B	216	Total	C	N	O	S	0	0
			1739	1099	308	322	10		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	220	GLY	-	expression tag	UNP Q8WYQ5
C	221	SER	-	expression tag	UNP Q8WYQ5
C	222	GLY	-	expression tag	UNP Q8WYQ5
B	220	GLY	-	expression tag	UNP Q8WYQ5
B	221	SER	-	expression tag	UNP Q8WYQ5
B	222	GLY	-	expression tag	UNP Q8WYQ5

- Molecule 3 is a RNA chain called Pri-miR-16-2 (78-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	78	Total	C	N	O	P	0	0
			1652	740	287	547	78		

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms		AltConf
4	A	2	Total 2	Zn 2	0

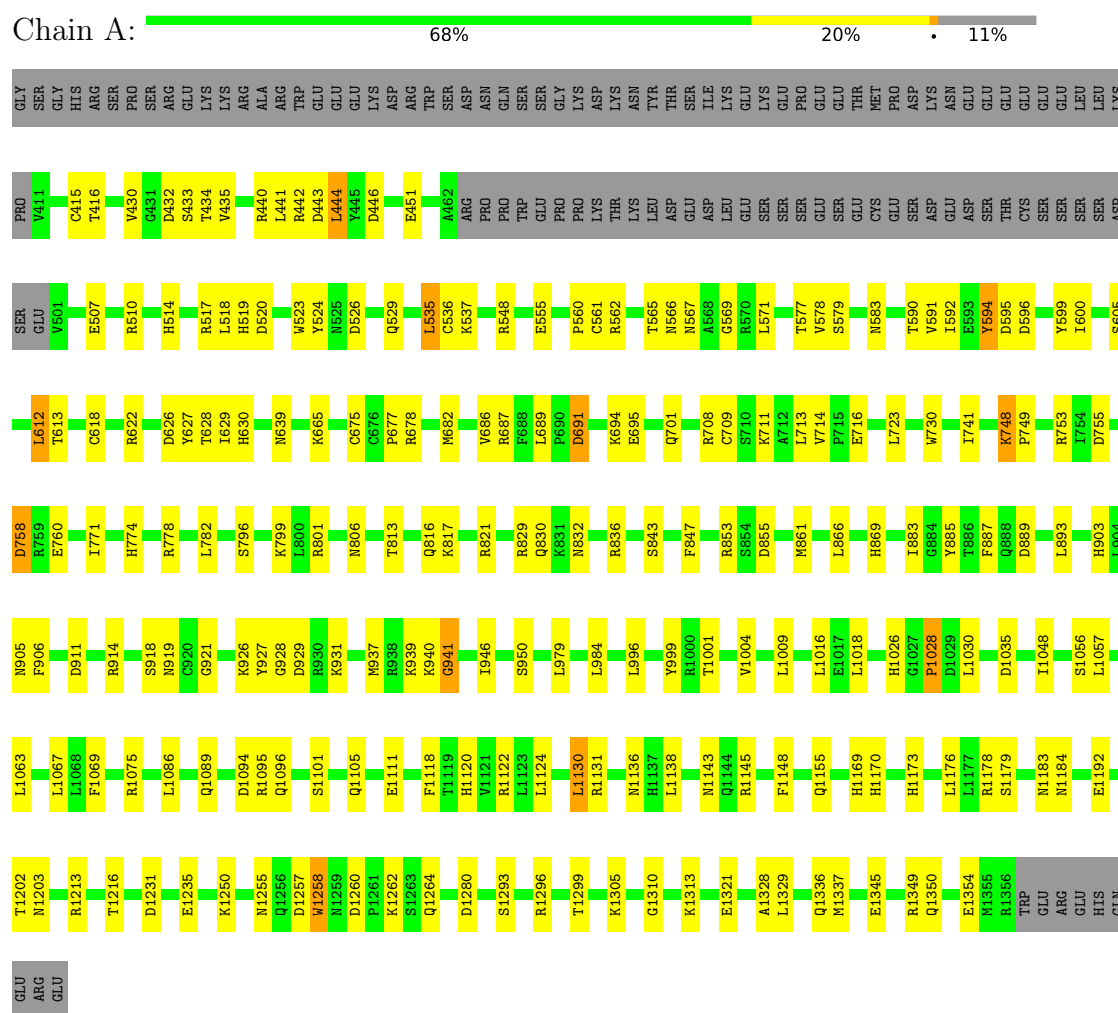
- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms		AltConf
5	A	1	Total 1	Ca 1	0

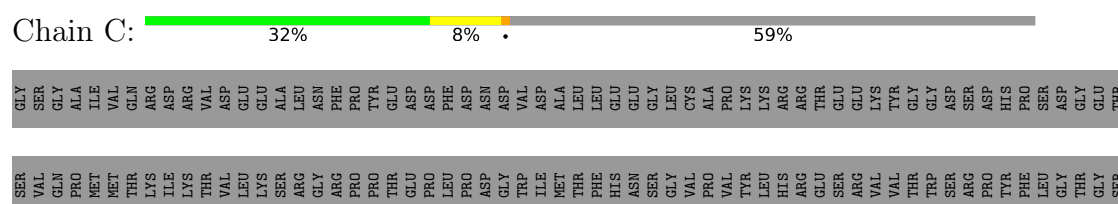
### 3 Residue-property plots

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: Ribonuclease 3



#### • Molecule 2: Microprocessor complex subunit DGCR8





## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	505640	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 ( $e^-/\text{\AA}^2$ )	46.8	Depositor
Minimum defocus (nm)	1300	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k 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: CA, ZN

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	0/7603	0.63	9/10279 (0.1%)
2	B	0.37	1/1769 (0.1%)	0.64	4/2369 (0.2%)
2	C	0.31	0/1780	0.58	2/2383 (0.1%)
3	D	0.77	0/1844	1.23	15/2866 (0.5%)
All	All	0.47	1/12996 (0.0%)	0.75	30/17897 (0.2%)

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
1	A	0	5
2	C	0	1
All	All	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	609	TYR	CB-CG	5.86	1.60	1.51

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	855	ASP	CB-CG-OD1	10.33	127.60	118.30
1	A	596	ASP	CB-CG-OD1	8.10	125.59	118.30
3	D	29	C	C2-N1-C1'	6.60	126.06	118.80
3	D	37	U	N1-C2-O2	6.49	127.34	122.80
3	D	15	C	N1-C2-O2	6.39	122.74	118.90



There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1028	PRO	Peptide
1	A	451	GLU	Peptide
1	A	714	VAL	Peptide
1	A	748	LYS	Peptide
1	A	941	GLY	Peptide

## 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	7419	0	7352	134	0
2	B	1739	0	1753	39	0
2	C	1750	0	1766	51	0
3	D	1652	0	837	28	0
4	A	2	0	0	0	0
5	A	1	0	0	0	0
All	All	12563	0	11708	236	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 236 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:C:609:TYR:HE2	2:C:674:GLN:OE1	1.16	1.25
2:C:609:TYR:HE2	2:C:674:GLN:CD	1.55	1.07
2:C:609:TYR:CE2	2:C:674:GLN:OE1	2.08	1.04
2:C:609:TYR:CE2	2:C:674:GLN:CD	2.32	1.03
2:C:609:TYR:CD1	2:C:620:PRO:CD	2.52	0.93

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	904/1016 (89%)	775 (86%)	127 (14%)	2 (0%)	49	83
2	B	206/532 (39%)	192 (93%)	14 (7%)	0	100	100
2	C	207/532 (39%)	194 (94%)	13 (6%)	0	100	100
All	All	1317/2080 (63%)	1161 (88%)	154 (12%)	2 (0%)	53	83

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	941	GLY
1	A	1028	PRO

### 5.3.2 Protein sidechains ⓘ

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	812/916 (89%)	806 (99%)	6 (1%)	85	93
2	B	193/475 (41%)	189 (98%)	4 (2%)	56	81
2	C	194/475 (41%)	193 (100%)	1 (0%)	90	96
All	All	1199/1866 (64%)	1188 (99%)	11 (1%)	82	91

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	906	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1258	TRP
2	B	630	ARG
1	A	691	ASP
2	B	610	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1336	GLN
2	C	531	ASN
2	B	622	GLN
1	A	1340	GLN
2	C	495	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	D	76/105 (72%)	18 (23%)	0

5 of 18 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	D	7	U
3	D	8	A
3	D	9	C
3	D	13	U
3	D	23	G

There are no RNA pucker outliers to report.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.