



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 06:31 am BST

PDB ID : 5B16  
Title : X-ray structure of DROSHA in complex with the C-terminal tail of DGCR8.  
Authors : Kwon, S.C.; Nguyen, T.A.; Choi, Y.G.; Jo, M.H.; Hohng, S.; Kim, V.N.; Woo, J.S.  
Deposited on : 2015-11-23  
Resolution : 3.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB 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/XrayValidationReportHelp>

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  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

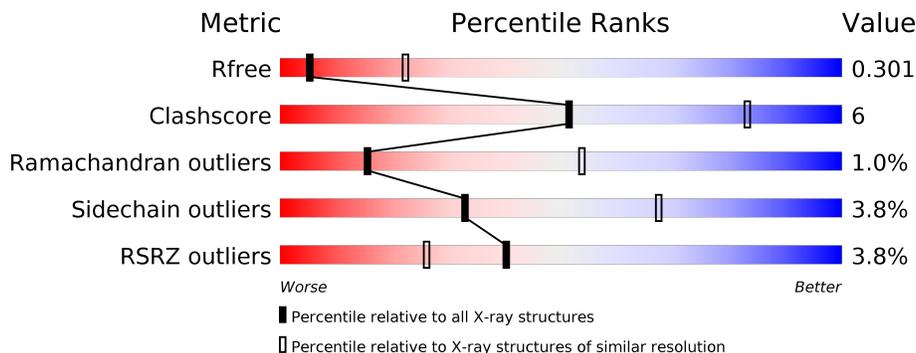
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	986	 3% 60% 12% 27%
2	B	39	 3% 59% 5% 36%
2	C	39	 56% 41%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 6159 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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, DRISHA, Ribonuclease 3, DRISHA, Ribonuclease 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	722	5754	3665	1007	1041	41	0	0	0

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	388	MET	-	initiating methionine	UNP Q9NRR4
A	389	ALA	-	expression tag	UNP Q9NRR4
A	390	LYS	-	expression tag	UNP Q9NRR4
A	391	GLU	-	expression tag	UNP Q9NRR4
A	392	PRO	-	expression tag	UNP Q9NRR4
A	393	GLU	-	expression tag	UNP Q9NRR4
A	394	GLU	-	expression tag	UNP Q9NRR4
A	395	THR	-	expression tag	UNP Q9NRR4
A	396	MET	-	expression tag	UNP Q9NRR4
A	397	PRO	-	expression tag	UNP Q9NRR4
A	398	ASP	-	expression tag	UNP Q9NRR4
A	399	LYS	-	expression tag	UNP Q9NRR4
A	400	ASN	-	expression tag	UNP Q9NRR4
A	401	GLU	-	expression tag	UNP Q9NRR4
A	402	GLU	-	expression tag	UNP Q9NRR4
A	403	GLU	-	expression tag	UNP Q9NRR4
A	404	GLU	-	expression tag	UNP Q9NRR4
A	405	GLU	-	expression tag	UNP Q9NRR4
A	406	GLU	-	expression tag	UNP Q9NRR4
A	407	LEU	-	expression tag	UNP Q9NRR4
A	408	LEU	-	expression tag	UNP Q9NRR4
A	409	LYS	-	expression tag	UNP Q9NRR4
A	410	PRO	-	expression tag	UNP Q9NRR4
A	1045	GLN	GLU	engineered mutation	UNP Q9NRR4
A	1222	GLN	GLU	engineered mutation	UNP Q9NRR4
A	1366	SER	-	expression tag	UNP Q9NRR4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	1367	ARG	-	expression tag	UNP Q9NRR4
A	1368	LEU	-	expression tag	UNP Q9NRR4
A	1369	GLU	-	expression tag	UNP Q9NRR4
A	1370	VAL	-	expression tag	UNP Q9NRR4
A	1371	LEU	-	expression tag	UNP Q9NRR4
A	1372	PHE	-	expression tag	UNP Q9NRR4
A	1373	GLN	-	expression tag	UNP Q9NRR4

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	25	211	128	40	41	2	0	0	0
2	C	23	192	117	35	39	1	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	726	MET	-	initiating methionine	UNP Q8WYQ5
B	727	ALA	-	expression tag	UNP Q8WYQ5
B	751	HIS	-	expression tag	UNP Q8WYQ5
B	752	GLY	-	expression tag	UNP Q8WYQ5
B	753	GLY	-	expression tag	UNP Q8WYQ5
B	754	SER	-	expression tag	UNP Q8WYQ5
B	755	ARG	-	expression tag	UNP Q8WYQ5
B	756	GLY	-	expression tag	UNP Q8WYQ5
B	757	ASP	-	expression tag	UNP Q8WYQ5
B	758	MET	-	expression tag	UNP Q8WYQ5
B	759	LEU	-	expression tag	UNP Q8WYQ5
B	760	GLU	-	expression tag	UNP Q8WYQ5
B	761	VAL	-	expression tag	UNP Q8WYQ5
B	762	LEU	-	expression tag	UNP Q8WYQ5
B	763	PHE	-	expression tag	UNP Q8WYQ5
B	764	GLN	-	expression tag	UNP Q8WYQ5
C	726	MET	-	initiating methionine	UNP Q8WYQ5
C	727	ALA	-	expression tag	UNP Q8WYQ5
C	751	HIS	-	expression tag	UNP Q8WYQ5
C	752	GLY	-	expression tag	UNP Q8WYQ5
C	753	GLY	-	expression tag	UNP Q8WYQ5
C	754	SER	-	expression tag	UNP Q8WYQ5
C	755	ARG	-	expression tag	UNP Q8WYQ5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	756	GLY	-	expression tag	UNP Q8WYQ5
C	757	ASP	-	expression tag	UNP Q8WYQ5
C	758	MET	-	expression tag	UNP Q8WYQ5
C	759	LEU	-	expression tag	UNP Q8WYQ5
C	760	GLU	-	expression tag	UNP Q8WYQ5
C	761	VAL	-	expression tag	UNP Q8WYQ5
C	762	LEU	-	expression tag	UNP Q8WYQ5
C	763	PHE	-	expression tag	UNP Q8WYQ5
C	764	GLN	-	expression tag	UNP Q8WYQ5

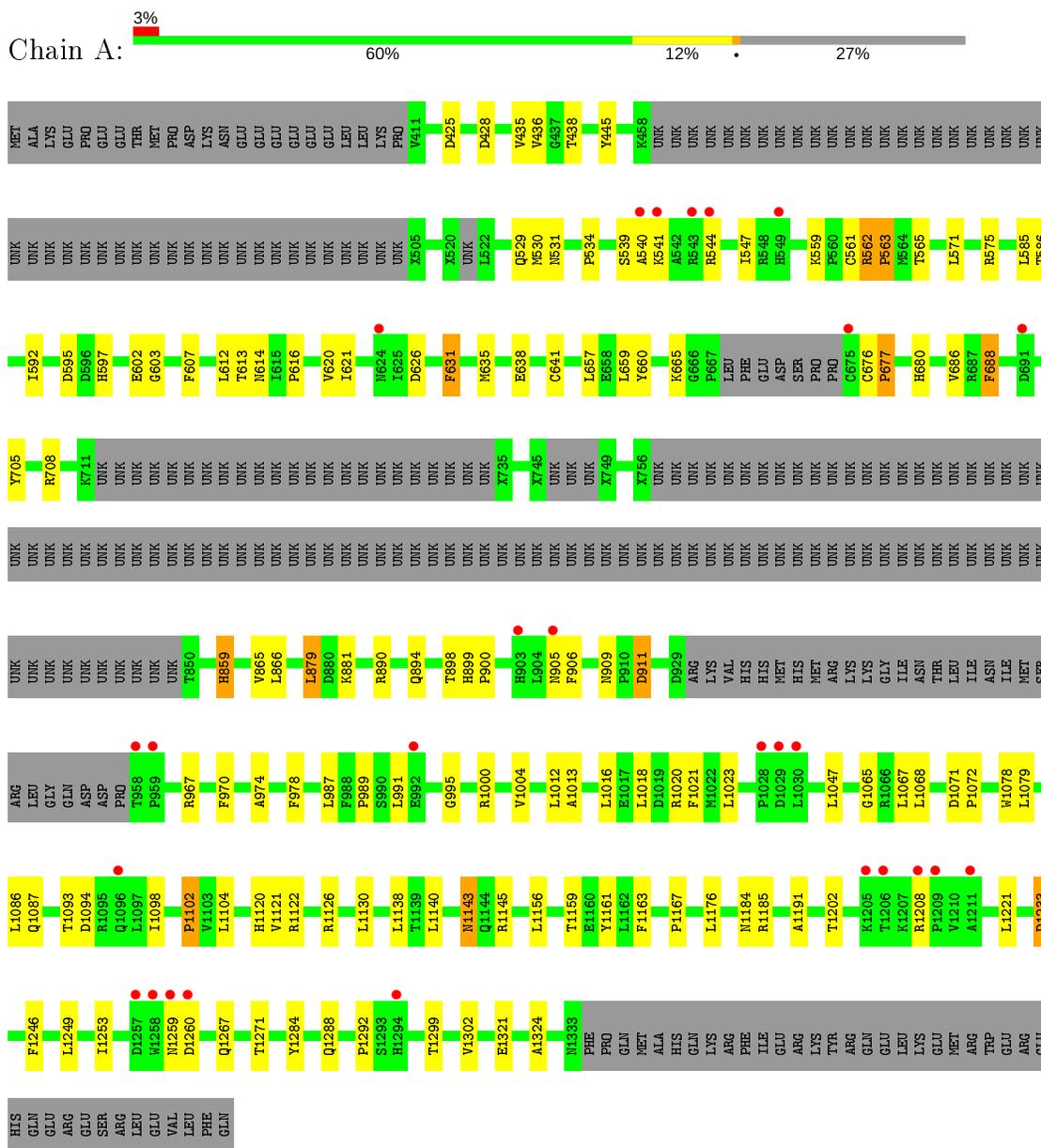
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Zn	0	0
			2	2		

### 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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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,DROSHA,Ribonuclease 3,DROSHA,Ribonuclease 3



- Molecule 2: Microprocessor complex subunit DGCR8



- Molecule 2: Microprocessor complex subunit DGCR8



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.25Å 118.14Å 122.30Å 90.00° 102.07° 90.00°	Depositor
Resolution (Å)	19.93 – 3.20 19.93 – 3.20	Depositor EDS
% Data completeness (in resolution range)	67.3 (19.93-3.20) 80.7 (19.93-3.20)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.29 (at 3.22Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.267 , 0.300 0.271 , 0.301	Depositor DCC
$R_{free}$ test set	2000 reflections (8.81%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.0	Xtrriage
Anisotropy	0.104	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 13.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.79	EDS
Total number of atoms	6159	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.46% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 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:  
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	0.21	0/5718	0.36	0/7739
2	B	0.20	0/211	0.32	0/278
2	C	0.19	0/192	0.35	0/254
All	All	0.21	0/6121	0.36	0/8271

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 [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	5754	0	5527	67	0
2	B	211	0	216	1	0
2	C	192	0	194	2	0
3	A	2	0	0	0	0
All	All	6159	0	5937	68	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 6.

The worst 5 of 68 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:1126:ARG:NH1	1:A:1140:LEU:O	2.23	0.71
1:A:1288:GLN:HB2	1:A:1299:THR:HB	1.81	0.63
1:A:621:ILE:HG12	1:A:626:ASP:HA	1.82	0.61
1:A:974:ALA:HB1	1:A:1156:LEU:HA	1.84	0.59
1:A:436:VAL:HA	1:A:565:THR:HG23	1.84	0.58

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 X-ray entries followed by that with respect to entries of similar resolution.

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	677/986 (69%)	601 (89%)	69 (10%)	7 (1%)	15	54
2	B	23/39 (59%)	23 (100%)	0	0	100	100
2	C	21/39 (54%)	21 (100%)	0	0	100	100
All	All	721/1064 (68%)	645 (90%)	69 (10%)	7 (1%)	15	54

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	561	CYS
1	A	613	THR
1	A	563	PRO
1	A	1259	ASN
1	A	677	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 X-ray entries followed by that with respect to entries of similar resolution.

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	612/706 (87%)	587 (96%)	25 (4%)	30	66
2	B	23/34 (68%)	23 (100%)	0	100	100
2	C	21/34 (62%)	21 (100%)	0	100	100
All	All	656/774 (85%)	631 (96%)	25 (4%)	33	67

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	866	LEU
1	A	881	LYS
1	A	1233	ASP
1	A	879	LEU
1	A	911	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 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](#)

Of 2 ligands modelled in this entry, 2 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.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ > 2	OWAB(Å <sup>2</sup> )	Q < 0.9
1	A	687/986 (69%)	0.03	27 (3%) 39 25	25, 54, 104, 166	0
2	B	25/39 (64%)	0.16	1 (4%) 38 25	58, 77, 100, 111	0
2	C	23/39 (58%)	0.25	0 100 100	52, 70, 99, 111	0
All	All	735/1064 (69%)	0.04	28 (3%) 40 26	25, 56, 104, 166	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1029	ASP	8.9
1	A	1257	ASP	6.2
1	A	1030	LEU	6.1
1	A	675	CYS	5.2
1	A	1259	ASN	3.8

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	ZN	A	1402	1/1	0.96	0.08	79,79,79,79	0
3	ZN	A	1401	1/1	0.99	0.06	31,31,31,31	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.