



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 02:10 am BST

PDB ID : 5EHK  
Title : Crystal structure of tRNA dependent lantibiotic dehydratase MibB from *Microbispora* sp. 107891  
Authors : Hao, Y.; Nair, S.K.  
Deposited on : 2015-10-28  
Resolution : 2.71 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

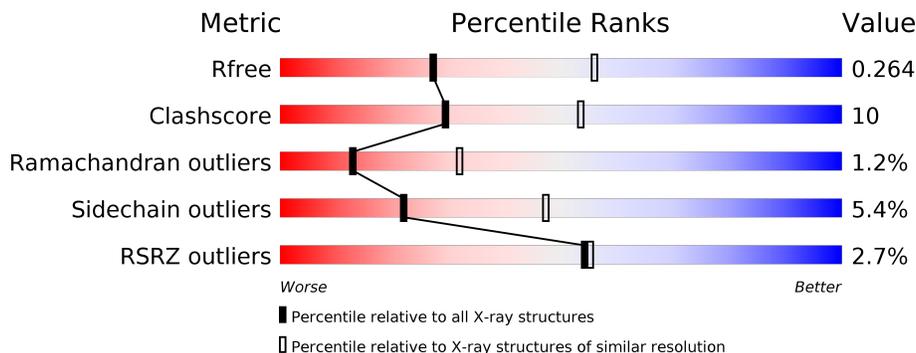
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.71 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	1115	<p>2% 70% 20% • 7%</p>
1	B	1115	<p>3% 72% 19% • 7%</p>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 16566 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 Lantibiotic dehydratase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1036	8044	5033	1533	1459	19	0	0	0
1	B	1036	8044	5033	1533	1459	19	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	230	Total 230	O 230	0	0
2	B	248	Total 248	O 248	0	0





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.14Å 151.11Å 205.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.24 – 2.71 45.24 – 2.71	Depositor EDS
% Data completeness (in resolution range)	96.1 (45.24-2.71) 96.1 (45.24-2.71)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.05 (at 2.73Å)	Xtrriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, $R_{free}$	0.204 , 0.264 0.204 , 0.264	Depositor DCC
$R_{free}$ test set	3537 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.3	Xtrriage
Anisotropy	0.129	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 41.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	16566	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.30% 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](#)

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	0.46	0/8223	0.67	3/11182 (0.0%)
1	B	0.47	0/8223	0.67	0/11182
All	All	0.47	0/16446	0.67	3/22364 (0.0%)

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	2
1	B	0	1
All	All	0	3

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	859	ARG	NE-CZ-NH2	-5.80	117.40	120.30
1	A	859	ARG	NE-CZ-NH1	5.74	123.17	120.30
1	A	1089	GLY	N-CA-C	-5.29	99.88	113.10

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1088	GLY	Peptide
1	A	780	PHE	Peptide
1	B	780	PHE	Peptide

## 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	8044	0	8012	156	1
1	B	8044	0	8012	158	1
2	A	230	0	0	26	0
2	B	248	0	0	31	0
All	All	16566	0	16024	306	1

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 306 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:B:541:PHE:O	2:B:1201:HOH:O	1.80	0.98
1:B:808:VAL:O	2:B:1202:HOH:O	1.84	0.93
1:A:688:LEU:O	2:A:1201:HOH:O	1.86	0.93
1:B:220:LYS:O	2:B:1203:HOH:O	1.85	0.93
1:B:786:GLU:OE2	2:B:1204:HOH:O	1.85	0.93

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1060:GLU:OE2	1:B:1053:ARG:NH2[4_445]	2.18	0.02

## 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	1024/1115 (92%)	963 (94%)	50 (5%)	11 (1%)	14	34
1	B	1024/1115 (92%)	963 (94%)	48 (5%)	13 (1%)	12	30
All	All	2048/2230 (92%)	1926 (94%)	98 (5%)	24 (1%)	13	32

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	116	GLY
1	A	427	LEU
1	A	1113	ARG
1	B	427	LEU
1	B	684	GLY

### 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	812/868 (94%)	760 (94%)	52 (6%)	17	39
1	B	812/868 (94%)	777 (96%)	35 (4%)	29	57
All	All	1624/1736 (94%)	1537 (95%)	87 (5%)	22	47

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	975	PRO
1	A	1072	LEU
1	B	1047	ARG
1	A	987	ASP
1	A	1054	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	183	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	832	GLN
1	A	1019	GLN
1	A	1080	HIS

### 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](#)

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.

## 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	1036/1115 (92%)	-0.09	26 (2%) 57 59	6, 32, 68, 108	0
1	B	1036/1115 (92%)	-0.10	30 (2%) 51 52	6, 31, 74, 106	0
All	All	2072/2230 (92%)	-0.09	56 (2%) 54 55	6, 32, 71, 108	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1114	HIS	8.9
1	B	118	LEU	7.9
1	A	707	PRO	5.5
1	B	742	ASP	5.4
1	A	1113	ARG	5.0

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.