

# wwPDB X-ray Structure Validation Summary Report (i)

#### May 21, 2020 - 01:52 am BST

PDB ID	:	5KTL
Title	:	Dihydrodipicolinate synthase from the industrial and evolutionarily important
		cyanobacteria Anabaena variabilis.
Authors	:	Christensen, J.B.; Panjikar, S.; Perugini, M.
Deposited on	:	2016-07-11
Resolution	:	1.92  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 A user guide is available at https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

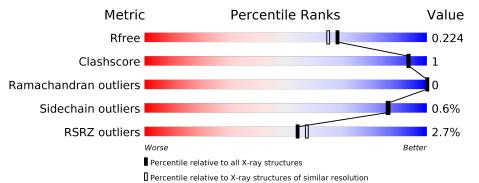
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
$\operatorname{Refmac}$	:	5.8.0158
CCP4	:	7.0.044   (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber $(2001)$
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X-RAY DIFFRACTION

The reported resolution of this entry is 1.92 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R <sub>free</sub>	130704	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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 >=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 <=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	А	330	86%	•	11%
1	В	330	2% 86%	•	11%



# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4699 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.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	Δ	295	Total	С	Ν	Ο	S	0	0	0
		290	2205	1401	373	425	6	0	0	0
1	В	295	Total	С	Ν	Ο	S	0	1	Ο
		290	2208	1403	373	426	6			0

• Molecule 1 is a protein called 4-hydroxy-tetrahydrodipicolinate synthase.

Chain	Residue	Modelled	Actual	Comment	Reference
А	-35	MET	-	initiating methionine	UNP Q3M723
А	-34	ARG	-	expression tag	UNP Q3M723
А	-33	GLY	-	expression tag	UNP Q3M723
А	-32	SER	-	expression tag	UNP Q3M723
A	-31	HIS	-	expression tag	UNP Q3M723
А	-30	HIS	-	expression tag	UNP Q3M723
А	-29	HIS	-	expression tag	UNP Q3M723
A	-28	HIS	-	expression tag	UNP Q3M723
А	-27	HIS	-	expression tag	UNP Q3M723
А	-26	HIS	-	expression tag	UNP Q3M723
А	-25	GLY	-	expression tag	UNP Q3M723
А	-24	MET	-	expression tag	UNP Q3M723
А	-23	ALA	-	expression tag	UNP Q3M723
А	-22	SER	-	expression tag	UNP Q3M723
A	-21	MET	-	expression tag	UNP Q3M723
A	-20	THR	-	expression tag	UNP Q3M723
A	-19	GLY	-	expression tag	UNP Q3M723
А	-18	GLY	-	expression tag	UNP Q3M723
А	-17	GLN	-	expression tag	UNP Q3M723
А	-16	GLN	-	expression tag	UNP Q3M723
А	-15	MET	-	expression tag	UNP Q3M723
А	-14	GLY	-	expression tag	UNP Q3M723
А	-13	ARG	-	expression tag	UNP Q3M723
А	-12	ASP	-	expression tag	UNP Q3M723
А	-11	LEU	_	expression tag	UNP Q3M723

There are 74 discrepancies between the modelled and reference sequences:

Continued on next page...



5KTL
------

	Continued from previous page							
Chain	Residue	Modelled	Actual	Comment	Reference			
А	-10	TYR	-	expression tag	UNP Q3M723			
A	-9	ASP	-	expression tag	UNP Q3M723			
А	-8	ASP	-	expression tag	UNP Q3M723			
А	-7	ASP	-	expression tag	UNP Q3M723			
А	-6	ASP	-	expression tag	UNP Q3M723			
А	-5	LYS	-	expression tag	UNP Q3M723			
А	-4	ASP	-	expression tag	UNP Q3M723			
А	-3	ARG	-	expression tag	UNP Q3M723			
А	-2	TRP	-	expression tag	UNP Q3M723			
А	-1	GLY	-	expression tag	UNP Q3M723			
А	0	SER	-	expression tag	UNP Q3M723			
A	1	VAL	-	expression tag	UNP Q3M723			
В	-35	MET	-	initiating methionine	UNP Q3M723			
В	-34	ARG	-	expression tag	UNP Q3M723			
В	-33	GLY	-	expression tag	UNP Q3M723			
В	-32	SER	-	expression tag	UNP Q3M723			
В	-31	HIS	-	expression tag	UNP Q3M723			
В	-30	HIS	-	expression tag	UNP Q3M723			
В	-29	HIS	-	expression tag	UNP Q3M723			
В	-28	HIS	-	expression tag	UNP Q3M723			
В	-27	HIS	-	expression tag	UNP Q3M723			
В	-26	HIS	-	expression tag	UNP Q3M723			
В	-25	GLY	-	expression tag	UNP Q3M723			
В	-24	MET	-	expression tag	UNP Q3M723			
В	-23	ALA	-	expression tag	UNP Q3M723			
В	-22	SER	-	expression tag	UNP Q3M723			
В	-21	MET	-	expression tag	UNP Q3M723			
В	-20	THR	-	expression tag	UNP Q3M723			
В	-19	GLY	-	expression tag	UNP Q3M723			
В	-18	GLY	-	expression tag	UNP Q3M723			
В	-17	GLN	-	expression tag	UNP Q3M723			
В	-16	GLN	-	expression tag	UNP Q3M723			
В	-15	MET	-	expression tag	UNP Q3M723			
В	-14	GLY	_	expression tag	UNP Q3M723			
В	-13	ARG	-	expression tag	UNP Q3M723			
В	-12	ASP	_	expression tag	UNP Q3M723			
В	-11	LEU	-	expression tag	UNP Q3M723			
В	-10	TYR	_	expression tag	UNP Q3M723			
В	-9	ASP	-	expression tag	UNP Q3M723			
В	-8	ASP	-	expression tag	UNP Q3M723			
В	-7	ASP	-	expression tag	UNP Q3M723			
В	-6	ASP	-	expression tag	UNP Q3M723			

Continued from previous page...

Continued on next page...



Continu	Continuea from previous page							
Chain	Residue	Modelled	Actual	Comment	Reference			
В	-5	LYS	-	expression tag	UNP Q3M723			
В	-4	ASP	-	expression tag	UNP Q3M723			
В	-3	ARG	-	expression tag	UNP Q3M723			
В	-2	TRP	-	expression tag	UNP Q3M723			
В	-1	GLY	-	expression tag	UNP Q3M723			
В	0	SER	-	expression tag	UNP Q3M723			
В	1	VAL	-	expression tag	UNP Q3M723			

Continued from previous page...

• Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	В	3	Total Mn 3 3	0	0
2	А	4	Total Mn 4 4	0	0

• Molecule 3 is water.

Ν	Aol	Chain	Residues	Atoms	ZeroOcc	AltConf
	3	А	137	Total O 137 137	0	0
	3	В	142	Total         O           142         142	0	0



#### Residue-property plots (i) 3

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.

- Chain A: 86% 11% • Molecule 1: 4-hydroxy-tetrahydrodipicolinate synthase Chain B: 86% 11%
- Molecule 1: 4-hydroxy-tetrahydrodipicolinate synthase





# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants	75.73Å $154.35$ Å $55.80$ Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	19.97 - 1.92	Depositor
Resolution (A)	19.97 - 1.92	EDS
% Data completeness	99.2 (19.97-1.92)	Depositor
(in resolution range)	99.3 (19.97 - 1.92)	EDS
R <sub>merge</sub>	(Not available)	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$0.94 (at 1.92 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
D D.	0.175 , $0.222$	Depositor
R, $R_{free}$	0.185 , $0.224$	DCC
$R_{free}$ test set	1016 reflections $(2.02\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	24.9	Xtriage
Anisotropy	0.137	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.42 , $47.2$	EDS
L-test for twinning <sup>2</sup>	$ \langle L  \rangle = 0.48, \langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4699	wwPDB-VP
Average B, all atoms $(Å^2)$	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.44% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: KPI, MN

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).

Mal	Mol Chain	Bond	lengths	Bond angles		
		RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.95	0/2230	0.88	2/3041~(0.1%)	
1	В	0.88	0/2236	0.91	3/3049~(0.1%)	
All	All	0.92	0/4466	0.90	5/6090~(0.1%)	

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
1	А	18	ASP	CB-CG-OD1	9.06	126.45	118.30
1	В	3	ASP	CB-CG-OD1	7.01	124.61	118.30
1	В	178	ARG	NE-CZ-NH2	-6.20	117.20	120.30
1	В	3	ASP	CB-CG-OD2	-5.90	112.99	118.30
1	А	18	ASP	CB-CG-OD2	-5.53	113.33	118.30

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	А	2205	0	2228	6	0
1	В	2208	0	2234	6	0
2	А	4	0	0	0	0
2	В	3	0	0	0	0

Continued on next page...



001100											
Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes					
3	А	137	0	0	1	0					
3	В	142	0	0	2	0					
All	All	4699	0	4462	12	0					

Continued from previous page...

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 1.

The worst 5 of 12 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:93:LYS:NZ	3:B:402:HOH:O	2.30	0.61
1:A:8:LEU:HD12	1:A:8:LEU:N	2.19	0.58
1:B:268:SER:HB2	1:B:275:ASP:OD1	2.07	0.55
1:B:70:ALA:HA	3:B:441:HOH:O	2.10	0.52
1:A:277:ASP:C	1:A:277:ASP:OD1	2.52	0.48

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	Perce	Percentiles	
1	А	292/330~(88%)	289~(99%)	3 (1%)	0	100	100	
1	В	293/330~(89%)	288~(98%)	5(2%)	0	100	100	
All	All	585/660~(89%)	577~(99%)	8 (1%)	0	100	100	

There are no Ramachandran outliers to report.



#### 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	А	235/265~(89%)	233~(99%)	2(1%)	78 78		
1	В	236/265~(89%)	234~(99%)	2(1%)	81 81		
All	All	471/530~(89%)	467~(99%)	4 (1%)	86 81		

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	67	GLN
1	А	279	GLU
1	В	233[A]	SER
1	В	233[B]	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	257	GLN
1	В	36	ASN
1	В	215	ASN

#### 5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the



Mol	Mol Type Chain		Chain Res Li		Bond lengths			Bond angles		
	Type	Chain	nes	Link	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2
1	KPI	В	164	1	10, 13, 14	1.04	0	$6,\!15,\!17$	1.88	2 (33%)
1	KPI	А	164	1	10, 13, 14	1.20	1 (10%)	$6,\!15,\!17$	2.22	2 (33%)

expected value. A bond length (or angle) with |Z| > 2 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	KPI	В	164	1	-	0/9/14/16	-
1	KPI	А	164	1	-	1/9/14/16	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
1	А	164	KPI	CB-CA	-2.47	1.50	1.53

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
1	А	164	KPI	C1-CX1-CX2	-4.07	113.42	117.92
1	А	164	KPI	CE-NZ-CX1	3.00	129.84	121.77
1	В	164	KPI	C1-CX1-CX2	-2.96	114.64	117.92
1	В	164	KPI	CE-NZ-CX1	2.44	128.32	121.77

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	А	164	KPI	O-C-CA-CB

There are no ring outliers.

No monomer is involved in short contacts.

### 5.5 Carbohydrates (i)

There are no carbohydrates in this entry.



#### 5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 7 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 (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	$OWAB(Å^2)$	Q<0.9
1	А	294/330~(89%)	-0.10	9 (3%) 49 52	17, 25, 46, 69	0
1	В	294/330 ( $89%$ )	-0.08	7 (2%) 59 62	18, 25, 48, 66	0
All	All	588/660~(89%)	-0.09	16 (2%) 54 57	17, 25, 47, 69	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	278	ALA	3.2
1	А	279	GLU	3.0
1	В	294	ILE	2.8
1	В	278	ALA	2.8
1	В	182	LYS	2.6

### 6.2 Non-standard residues in protein, DNA, RNA chains (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,  $95^{th}$  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	$\mathbf{B} ext{-factors}(\mathbf{\AA}^2)$	Q < 0.9
1	KPI	В	164	14/15	0.96	0.07	$16,\!21,\!26,\!27$	0
1	KPI	А	164	14/15	0.96	0.08	$16,\!19,\!26,\!26$	0

### 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,  $95^{th}$  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	$\mathbf{B} ext{-factors}(\mathbf{A}^2)$	$Q{<}0.9$
2	MN	А	304	1/1	0.92	0.12	$90,\!90,\!90,\!90$	0
2	MN	В	302	1/1	0.93	0.11	76, 76, 76, 76, 76	0
2	MN	А	303	1/1	0.96	0.11	$65,\!65,\!65,\!65$	0
2	MN	А	302	1/1	0.97	0.05	$60,\!60,\!60,\!60$	0
2	MN	В	303	1/1	0.97	0.09	$67,\!67,\!67,\!67$	0
2	MN	В	301	1/1	0.98	0.05	$59,\!59,\!59,\!59$	0
2	MN	А	301	1/1	1.00	0.06	34,34,34,34	0

# 6.5 Other polymers (i)

There are no such residues in this entry.

