



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 31, 2021 – 10:22 AM EDT

PDB ID : 6VVI
Title : Arabidopsis thaliana dihydrodipicolinate synthase isoform 1 (DHDPS1)
Authors : Lee, M.; Hall, C.J.
Deposited on : 2020-02-18
Resolution : 2.15 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.23.1
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) : 2.23.1

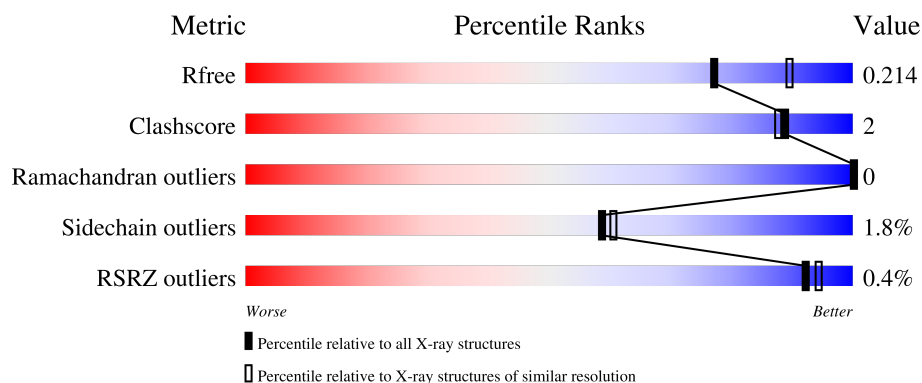
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2523 (2.16-2.12)
Clashscore	141614	2653 (2.16-2.12)
Ramachandran outliers	138981	2618 (2.16-2.12)
Sidechain outliers	138945	2617 (2.16-2.12)
RSRZ outliers	127900	2485 (2.16-2.12)

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 of 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 $\leq 5\%$. 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	AAA	321	<div> <div>%</div> <div> <div></div> <div>91%</div> <div>5%</div> <div>.</div> </div> </div>
1	BBB	321	<div> <div>93%</div> <div>.</div> <div>.</div> <div>.</div> </div>
1	CCC	321	<div> <div>90%</div> <div>6%</div> <div>.</div> <div>.</div> </div>
1	DDD	321	<div> <div>93%</div> <div>.</div> <div>.</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	SO4	DDD	404	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10206 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 4-hydroxy-tetrahydrodipicolinate synthase 1, chloroplastic.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	308	Total	C	N	O	S	0	1	0
			2401	1514	425	448	14			
1	BBB	312	Total	C	N	O	S	0	4	0
			2455	1545	436	460	14			
1	CCC	309	Total	C	N	O	S	0	2	0
			2413	1522	425	452	14			
1	DDD	311	Total	C	N	O	S	0	1	0
			2419	1524	426	455	14			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	45	GLY	-	expression tag	UNP Q9LZX6
AAA	46	PRO	-	expression tag	UNP Q9LZX6
AAA	47	GLY	-	expression tag	UNP Q9LZX6
AAA	48	SER	-	expression tag	UNP Q9LZX6
BBB	45	GLY	-	expression tag	UNP Q9LZX6
BBB	46	PRO	-	expression tag	UNP Q9LZX6
BBB	47	GLY	-	expression tag	UNP Q9LZX6
BBB	48	SER	-	expression tag	UNP Q9LZX6
CCC	45	GLY	-	expression tag	UNP Q9LZX6
CCC	46	PRO	-	expression tag	UNP Q9LZX6
CCC	47	GLY	-	expression tag	UNP Q9LZX6
CCC	48	SER	-	expression tag	UNP Q9LZX6
DDD	45	GLY	-	expression tag	UNP Q9LZX6
DDD	46	PRO	-	expression tag	UNP Q9LZX6
DDD	47	GLY	-	expression tag	UNP Q9LZX6
DDD	48	SER	-	expression tag	UNP Q9LZX6

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	AAA	1	Total	C	O	0	0
			6	3	3		
2	CCC	1	Total	C	O	0	0
			6	3	3		
2	DDD	1	Total	C	O	0	0
			6	3	3		
2	DDD	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	1	Total	Na	0	0
			1	1		
3	BBB	1	Total	Na	0	0
			1	1		
3	CCC	1	Total	Na	0	0
			1	1		
3	DDD	1	Total	Na	0	0
			1	1		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	BBB	1	Total	O	S	0	0
			5	4	1		
4	CCC	1	Total	O	S	0	0
			5	4	1		
4	CCC	1	Total	O	S	0	0
			5	4	1		
4	DDD	1	Total	O	S	0	0
			5	4	1		
4	DDD	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	85	Total	O	0	0
			85	85		
5	BBB	130	Total	O	0	0
			130	130		
5	CCC	107	Total	O	0	0
			107	107		
5	DDD	143	Total	O	0	0
			143	143		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

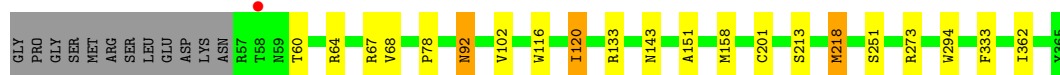
- Molecule 1: 4-hydroxy-tetrahydrodipicolinate synthase 1, chloroplastic



- Molecule 1: 4-hydroxy-tetrahydrodipicolinate synthase 1, chloroplastic



- Molecule 1: 4-hydroxy-tetrahydrodipicolinate synthase 1, chloroplastic



- Molecule 1: 4-hydroxy-tetrahydrodipicolinate synthase 1, chloroplastic



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	95.47Å 98.10Å 176.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.10 – 2.15 49.05 – 2.14	Depositor EDS
% Data completeness (in resolution range)	99.2 (49.10-2.15) 99.2 (49.05-2.14)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.29 (at 2.14Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.172 , 0.209 0.180 , 0.214	Depositor DCC
R_{free} test set	4371 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å ²)	38.7	Xtriage
Anisotropy	0.505	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 33.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.032 for k,h,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	10206	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: NA, SO4, GOL

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	AAA	0.64	0/2456	0.77	0/3332
1	BBB	0.64	0/2519	0.76	0/3416
1	CCC	0.65	0/2472	0.78	0/3355
1	DDD	0.65	0/2474	0.77	0/3358
All	All	0.65	0/9921	0.77	0/13461

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	AAA	2401	0	2368	6	0
1	BBB	2455	0	2423	7	0
1	CCC	2413	0	2375	17	0
1	DDD	2419	0	2376	7	0
2	AAA	6	0	8	1	0
2	CCC	6	0	8	3	0
2	DDD	12	0	16	0	0
3	AAA	1	0	0	0	0
3	BBB	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	CCC	1	0	0	0	0
3	DDD	1	0	0	0	0
4	BBB	5	0	0	0	0
4	CCC	10	0	0	1	0
4	DDD	10	0	0	2	0
5	AAA	85	0	0	0	0
5	BBB	130	0	0	2	0
5	CCC	107	0	0	1	0
5	DDD	143	0	0	1	0
All	All	10206	0	9574	36	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 2.

The worst 5 of 36 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:CCC:116:TRP:CZ2	1:CCC:120:ILE:HG23	2.13	0.83
1:BBB:227:ASN:ND2	1:BBB:251:SER:OG	2.20	0.75
1:DDD:227:ASN:ND2	1:DDD:251:SER:OG	2.22	0.73
1:DDD:213:SER:HA	1:DDD:218:MET:HG2	1.74	0.68
1:CCC:218:MET:O	1:CCC:218:MET:HG3	1.94	0.68

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 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	AAA	307/321 (96%)	301 (98%)	6 (2%)	0	100	100
1	BBB	314/321 (98%)	309 (98%)	5 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	CCC	309/321 (96%)	303 (98%)	6 (2%)	0	100	100
1	DDD	310/321 (97%)	306 (99%)	4 (1%)	0	100	100
All	All	1240/1284 (97%)	1219 (98%)	21 (2%)	0	100	100

There are no Ramachandran outliers to report.

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 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	AAA	261/272 (96%)	255 (98%)	6 (2%)	50	51
1	BBB	268/272 (98%)	263 (98%)	5 (2%)	57	59
1	CCC	263/272 (97%)	257 (98%)	6 (2%)	50	51
1	DDD	263/272 (97%)	261 (99%)	2 (1%)	81	85
All	All	1055/1088 (97%)	1036 (98%)	19 (2%)	59	60

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	201	CYS
1	DDD	201	CYS
1	DDD	301	PRO
1	CCC	251	SER
1	BBB	251	SER

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

5.3.3 RNA ⓘ

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 4 are monoatomic - leaving 9 for Mogul analysis.

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 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).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	SO4	CCC	403	-	4,4,4	0.39	0	6,6,6	0.07	0
2	GOL	CCC	401	-	5,5,5	0.16	0	5,5,5	0.36	0
4	SO4	DDD	405	-	4,4,4	0.48	0	6,6,6	0.14	0
2	GOL	DDD	402	-	5,5,5	0.13	0	5,5,5	0.40	0
4	SO4	DDD	404	-	4,4,4	0.41	0	6,6,6	0.13	0
4	SO4	BBB	402	-	4,4,4	0.37	0	6,6,6	0.08	0
4	SO4	CCC	404	-	4,4,4	0.46	0	6,6,6	0.11	0
2	GOL	AAA	401	-	5,5,5	0.13	0	5,5,5	0.38	0
2	GOL	DDD	401	-	5,5,5	0.26	0	5,5,5	0.60	0

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
2	GOL	CCC	401	-	-	2/4/4/4	-
2	GOL	AAA	401	-	-	2/4/4/4	-
2	GOL	DDD	401	-	-	3/4/4/4	-
2	GOL	DDD	402	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	CCC	401	GOL	O1-C1-C2-C3
2	DDD	401	GOL	C1-C2-C3-O3
2	DDD	402	GOL	O1-C1-C2-C3
2	DDD	402	GOL	C1-C2-C3-O3
2	DDD	401	GOL	O2-C2-C3-O3

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	CCC	401	GOL	3	0
4	DDD	404	SO4	2	0
4	CCC	404	SO4	1	0
2	AAA	401	GOL	1	0

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, 95th 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	AAA	308/321 (95%)	-0.20	2 (0%) 89 91	33, 45, 64, 88	0
1	BBB	312/321 (97%)	-0.40	1 (0%) 94 95	31, 39, 55, 98	0
1	CCC	309/321 (96%)	-0.23	1 (0%) 94 95	32, 43, 59, 85	0
1	DDD	311/321 (96%)	-0.32	1 (0%) 94 95	30, 40, 58, 82	0
All	All	1240/1284 (96%)	-0.28	5 (0%) 92 94	30, 41, 61, 98	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	339	GLU	3.1
1	CCC	58	THR	3.0
1	BBB	54	ASP	2.5
1	DDD	349	ARG	2.5
1	AAA	348	ASP	2.1

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 monosaccharides 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, 95th 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(Å ²)	Q<0.9
2	GOL	DDD	401	6/6	0.87	0.15	48,55,59,61	0
2	GOL	CCC	401	6/6	0.88	0.21	51,58,65,65	0
4	SO4	CCC	403	5/5	0.89	0.18	49,52,54,54	5
4	SO4	DDD	404	5/5	0.90	0.24	48,49,50,54	5
4	SO4	DDD	405	5/5	0.90	0.25	41,42,45,53	5
2	GOL	DDD	402	6/6	0.91	0.25	56,62,63,64	0
2	GOL	AAA	401	6/6	0.93	0.16	48,54,57,62	0
4	SO4	BBB	402	5/5	0.94	0.16	51,51,54,56	5
3	NA	AAA	402	1/1	0.94	0.18	40,40,40,40	0
4	SO4	CCC	404	5/5	0.96	0.09	39,41,42,43	5
3	NA	BBB	401	1/1	0.96	0.16	37,37,37,37	0
3	NA	CCC	402	1/1	0.96	0.17	39,39,39,39	0
3	NA	DDD	403	1/1	0.98	0.13	33,33,33,33	0

6.5 Other polymers [i](#)

There are no such residues in this entry.