



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 08:42 am BST

PDB ID : 6V0K  
Title : Crystal structure of *Penicillium verruculosum* copalyl diphosphate synthase (PvCPS) alpha prenyltransferase domain  
Authors : Christianson, D.W.; Ronnebaum, T.A.  
Deposited on : 2019-11-18  
Resolution : 2.41 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
Mogul	:	1.8.5 (274361), CSD as541be (2020)
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

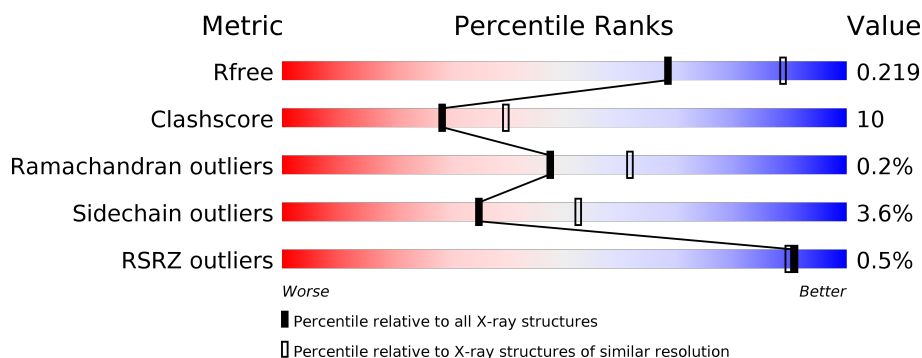
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.41 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	4647 (2.44-2.40)
Clashscore	141614	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)
RSRZ outliers	127900	4543 (2.44-2.40)

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	304	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: -10px; left: 0;">%</div> <div style="position: absolute; top: 0; left: 0; width: 100%; height: 100%; background: linear-gradient(to right, red, orange, yellow, green);"></div> <div style="position: absolute; bottom: -10px; left: 0;">84%</div> <div style="position: absolute; bottom: -10px; right: 0;">14%</div> <div style="position: absolute; bottom: -10px; right: 0;">..</div> </div> </div>
1	B	304	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: 0; left: 0; width: 100%; height: 100%; background: linear-gradient(to right, green, yellow, orange, red, grey);"></div> <div style="position: absolute; bottom: -10px; left: 0;">75%</div> <div style="position: absolute; bottom: -10px; right: 0;">20%</div> <div style="position: absolute; bottom: -10px; right: 0;">..</div> </div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5030 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 Terpene synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	302	Total	C	N	O	S	0	0	0
			2387	1517	404	453	13			
1	B	296	Total	C	N	O	S	0	0	0
			2355	1502	400	440	13			

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		

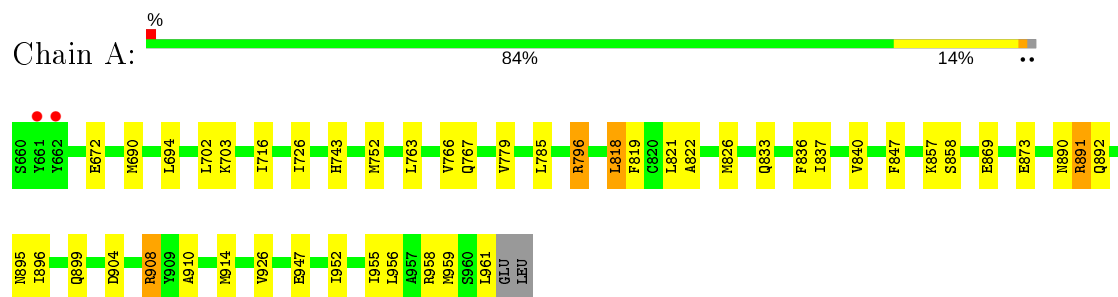
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	135	Total 135	O 135	0	0
3	B	141	Total 141	O 141	0	0

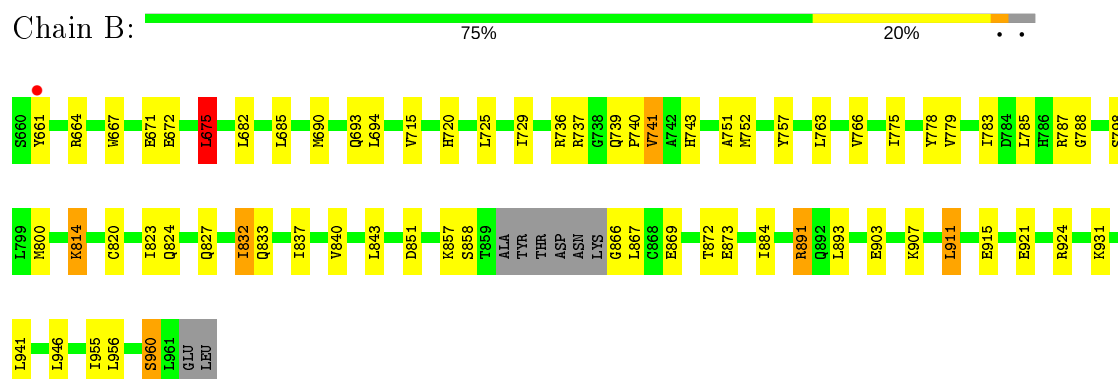
### 3 Residue-property plots [i](#)

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: Terpene synthase



#### • Molecule 1: Terpene synthase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	188.12Å 188.12Å 56.41Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	81.46 – 2.41 81.46 – 2.41	Depositor EDS
% Data completeness (in resolution range)	99.8 (81.46-2.41) 99.8 (81.46-2.41)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.45 (at 2.40Å)	Xtriage
Refinement program	PHENIX 1.13 _2998	Depositor
R, $R_{free}$	0.183 , 0.219 0.183 , 0.219	Depositor DCC
$R_{free}$ test set	2000 reflections (4.52%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.3	Xtriage
Anisotropy	0.442	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 47.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.051 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5030	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.94% 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: 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	A	0.41	0/2429	0.56	0/3290
1	B	0.43	0/2396	0.57	1/3238 (0.0%)
All	All	0.42	0/4825	0.57	1/6528 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	675	LEU	CA-CB-CG	-5.26	103.21	115.30

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	A	2387	0	2368	43	0
1	B	2355	0	2371	56	0
2	A	6	0	8	0	0
2	B	6	0	8	3	0
3	A	135	0	0	5	0
3	B	141	0	0	3	0
All	All	5030	0	4755	93	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 93 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:952:ILE:HD12	1:A:955:ILE:HD11	1.43	1.00
1:A:891:ARG:HH11	1:A:891:ARG:HG3	1.40	0.87
1:A:752:MET:HG3	1:B:752:MET:CE	2.09	0.82
1:B:891:ARG:HG3	1:B:891:ARG:HH11	1.50	0.76
1:B:921:GLU:HG2	1:B:924:ARG:HH21	1.51	0.75

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	300/304 (99%)	292 (97%)	8 (3%)	0	100	100
1	B	292/304 (96%)	285 (98%)	6 (2%)	1 (0%)	41	54
All	All	592/608 (97%)	577 (98%)	14 (2%)	1 (0%)	47	61

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	960	SER

### 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	261/272 (96%)	253 (97%)	8 (3%)	40	58
1	B	260/272 (96%)	249 (96%)	11 (4%)	30	46
All	All	521/544 (96%)	502 (96%)	19 (4%)	35	52

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

Mol	Chain	Res	Type
1	B	664	ARG
1	B	690	MET
1	B	851	ASP
1	A	961	LEU
1	B	891	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	743	HIS
1	B	789	GLN

### 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](#)

2 ligands 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 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
2	GOL	B	1001	-	5,5,5	0.96	0	5,5,5	1.27	1 (20%)
2	GOL	A	1001	-	5,5,5	1.03	0	5,5,5	0.92	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	B	1001	-	-	2/4/4/4	-
2	GOL	A	1001	-	-	2/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1001	GOL	C3-C2-C1	-2.31	102.70	111.70

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1001	GOL	C1-C2-C3-O3
2	B	1001	GOL	O2-C2-C3-O3
2	A	1001	GOL	C1-C2-C3-O3
2	A	1001	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1001	GOL	3	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, 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	302/304 (99%)	-0.60	2 (0%) 87 86	17, 27, 58, 74	0
1	B	296/304 (97%)	-0.55	1 (0%) 94 93	17, 28, 51, 70	0
All	All	598/608 (98%)	-0.58	3 (0%) 91 89	17, 27, 54, 74	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	661	TYR	3.0
1	A	662	TYR	2.6
1	B	661	TYR	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 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(Å <sup>2</sup> )	Q<0.9
2	GOL	B	1001	6/6	0.85	0.15	53,58,64,70	0
2	GOL	A	1001	6/6	0.88	0.14	48,52,53,59	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.