



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 26, 2023 – 10:28 PM JST

PDB ID : 8KBW  
Title : The crystal structure of syn-copalyl diphosphate synthase from *Oryza sativa*  
Authors : Ma, X.L.; Xu, H.F.; Jiang, T.  
Deposited on : 2023-08-04  
Resolution : 3.49 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
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) : 2.36

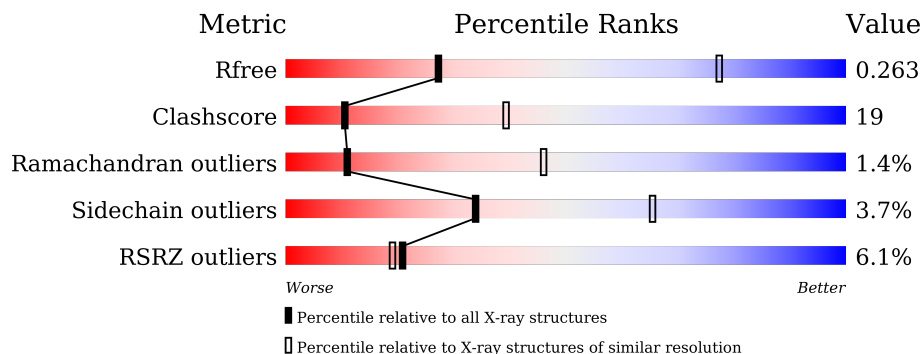
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.49 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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 of 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	775	 5% 51% 35% 11%
1	B	775	 5% 54% 31% 14%
1	C	775	 9% 51% 30% 16%
1	D	775	 8% 54% 33% 11%
1	E	775	 57% 30% 11%
1	F	775	 9% 52% 33% 13%

## 2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 32467 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 Syn-copalyl diphosphate synthase, chloroplastic.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	668	Total 5341	C 3391	N 916	O 1000	S 34	0	0	0
1	C	648	Total 5186	C 3297	N 892	O 962	S 35	0	1	0
1	E	688	Total 5504	C 3492	N 944	O 1033	S 35	0	0	0
1	D	688	Total 5504	C 3492	N 944	O 1033	S 35	0	0	0
1	F	677	Total 5421	C 3439	N 928	O 1019	S 35	0	0	0
1	A	689	Total 5511	C 3496	N 945	O 1035	S 35	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

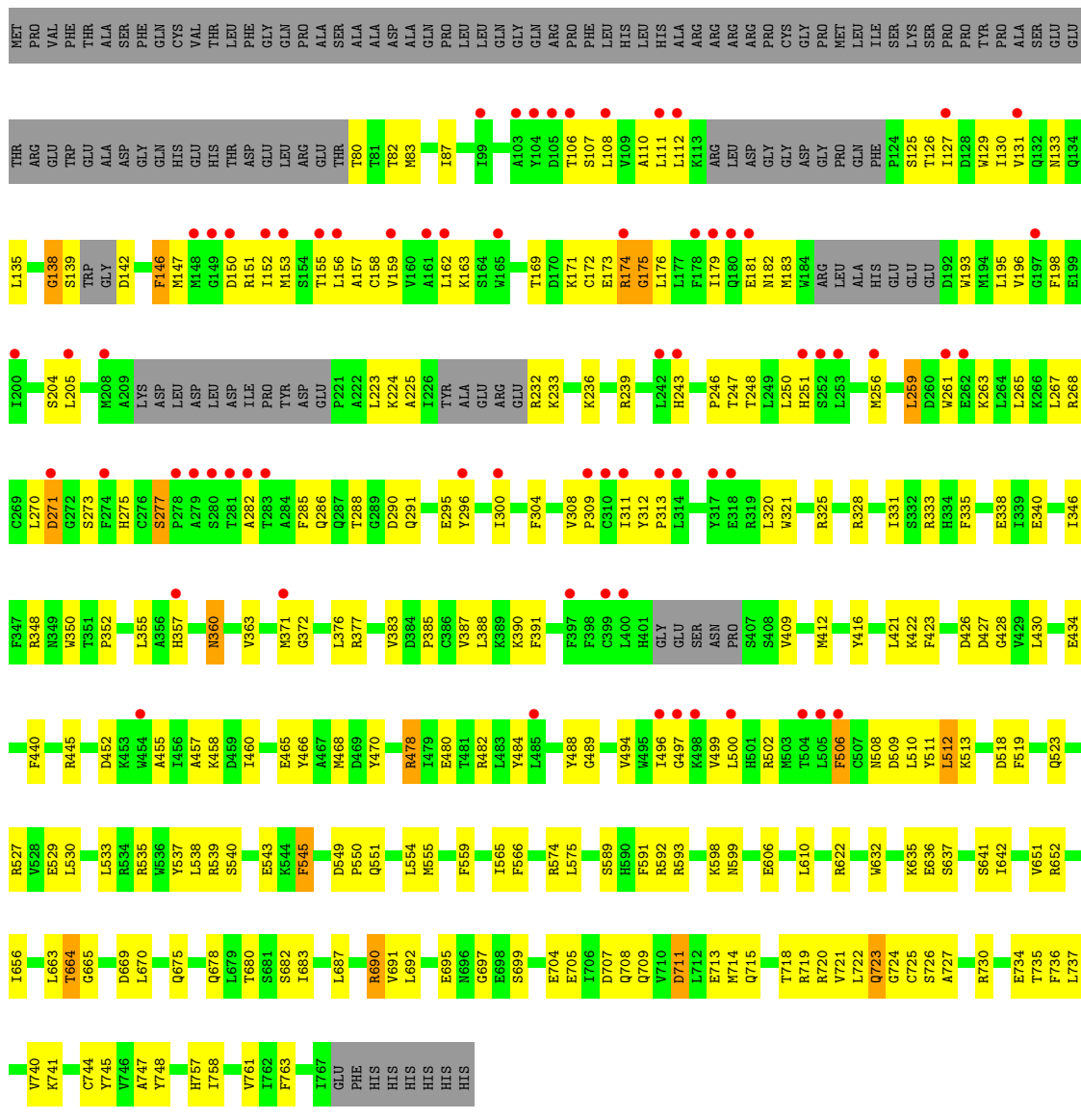
Chain	Residue	Modelled	Actual	Comment	Reference
B	768	GLU	-	expression tag	UNP Q0JF02
B	769	PHE	-	expression tag	UNP Q0JF02
B	770	HIS	-	expression tag	UNP Q0JF02
B	771	HIS	-	expression tag	UNP Q0JF02
B	772	HIS	-	expression tag	UNP Q0JF02
B	773	HIS	-	expression tag	UNP Q0JF02
B	774	HIS	-	expression tag	UNP Q0JF02
B	775	HIS	-	expression tag	UNP Q0JF02
C	768	GLU	-	expression tag	UNP Q0JF02
C	769	PHE	-	expression tag	UNP Q0JF02
C	770	HIS	-	expression tag	UNP Q0JF02
C	771	HIS	-	expression tag	UNP Q0JF02
C	772	HIS	-	expression tag	UNP Q0JF02
C	773	HIS	-	expression tag	UNP Q0JF02
C	774	HIS	-	expression tag	UNP Q0JF02
C	775	HIS	-	expression tag	UNP Q0JF02
E	768	GLU	-	expression tag	UNP Q0JF02

*Continued on next page...*

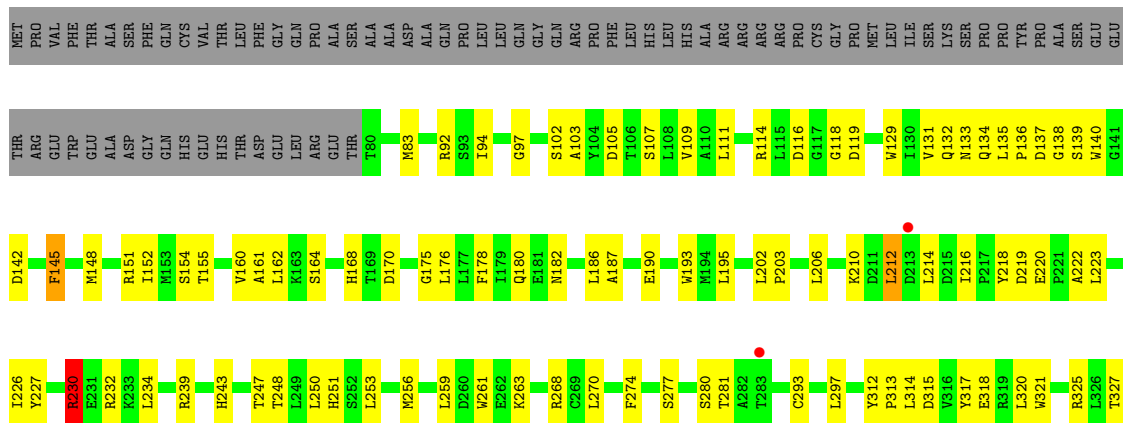
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	769	PHE	-	expression tag	UNP Q0JF02
E	770	HIS	-	expression tag	UNP Q0JF02
E	771	HIS	-	expression tag	UNP Q0JF02
E	772	HIS	-	expression tag	UNP Q0JF02
E	773	HIS	-	expression tag	UNP Q0JF02
E	774	HIS	-	expression tag	UNP Q0JF02
E	775	HIS	-	expression tag	UNP Q0JF02
D	768	GLU	-	expression tag	UNP Q0JF02
D	769	PHE	-	expression tag	UNP Q0JF02
D	770	HIS	-	expression tag	UNP Q0JF02
D	771	HIS	-	expression tag	UNP Q0JF02
D	772	HIS	-	expression tag	UNP Q0JF02
D	773	HIS	-	expression tag	UNP Q0JF02
D	774	HIS	-	expression tag	UNP Q0JF02
D	775	HIS	-	expression tag	UNP Q0JF02
F	768	GLU	-	expression tag	UNP Q0JF02
F	769	PHE	-	expression tag	UNP Q0JF02
F	770	HIS	-	expression tag	UNP Q0JF02
F	771	HIS	-	expression tag	UNP Q0JF02
F	772	HIS	-	expression tag	UNP Q0JF02
F	773	HIS	-	expression tag	UNP Q0JF02
F	774	HIS	-	expression tag	UNP Q0JF02
F	775	HIS	-	expression tag	UNP Q0JF02
A	768	GLU	-	expression tag	UNP Q0JF02
A	769	PHE	-	expression tag	UNP Q0JF02
A	770	HIS	-	expression tag	UNP Q0JF02
A	771	HIS	-	expression tag	UNP Q0JF02
A	772	HIS	-	expression tag	UNP Q0JF02
A	773	HIS	-	expression tag	UNP Q0JF02
A	774	HIS	-	expression tag	UNP Q0JF02
A	775	HIS	-	expression tag	UNP Q0JF02





● Molecule 1: Syn-copalyl diphosphate synthase, chloroplactic











## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	129.70Å 174.34Å 296.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	97.62 – 3.49 98.19 – 3.49	Depositor EDS
% Data completeness (in resolution range)	96.5 (97.62-3.49) 95.4 (98.19-3.49)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.63 (at 3.49Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.201 , 0.263 0.201 , 0.263	Depositor DCC
$R_{free}$ test set	2000 reflections (2.33%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	86.5	Xtrriage
Anisotropy	0.366	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 56.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	32467	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.38% 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	0.63	0/5634	0.84	4/7625 (0.1%)
1	B	0.54	0/5455	0.74	1/7375 (0.0%)
1	C	0.54	0/5297	0.75	1/7155 (0.0%)
1	D	0.51	0/5627	0.70	1/7615 (0.0%)
1	E	0.57	2/5627 (0.0%)	0.77	4/7615 (0.1%)
1	F	0.49	0/5540	0.75	3/7496 (0.0%)
All	All	0.55	2/33180 (0.0%)	0.76	14/44881 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	744	CYS	CB-SG	-6.91	1.70	1.82
1	E	698	GLU	CB-CG	5.77	1.63	1.52

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	177	LEU	CA-CB-CG	6.72	130.76	115.30
1	E	234	LEU	CA-CB-CG	6.52	130.29	115.30
1	F	249	LEU	CA-CB-CG	6.07	129.27	115.30
1	E	629	LEU	CA-CB-CG	-6.05	101.38	115.30
1	F	530	LEU	CA-CB-CG	5.90	128.87	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	5511	0	5420	228	0
1	B	5341	0	5268	196	0
1	C	5186	0	5140	199	0
1	D	5504	0	5413	199	0
1	E	5504	0	5413	181	0
1	F	5421	0	5323	224	0
All	All	32467	0	31977	1193	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 19.

The worst 5 of 1193 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:E:608:ILE:HD11	1:E:621:LEU:HD22	1.34	1.08
1:A:156:LEU:HD11	1:A:205:LEU:HG	1.38	1.04
1:F:460:ILE:HG23	1:F:461:PRO:HD3	1.49	0.94
1:B:606:GLU:O	1:B:608:ILE:N	2.01	0.92
1:B:156:LEU:HD11	1:B:205:LEU:HG	1.52	0.91

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	687/775 (89%)	621 (90%)	56 (8%)	10 (2%)	10	45
1	B	658/775 (85%)	611 (93%)	37 (6%)	10 (2%)	10	45
1	C	635/775 (82%)	571 (90%)	52 (8%)	12 (2%)	8	40
1	D	686/775 (88%)	637 (93%)	41 (6%)	8 (1%)	13	50
1	E	686/775 (88%)	630 (92%)	49 (7%)	7 (1%)	15	54

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	671/775 (87%)	605 (90%)	57 (8%)	9 (1%)	12	48
All	All	4023/4650 (86%)	3675 (91%)	292 (7%)	56 (1%)	11	46

5 of 56 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	403	GLU
1	B	607	LEU
1	B	697	GLY
1	C	183	MET
1	C	664	THR

### 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	598/672 (89%)	571 (96%)	27 (4%)	27	61
1	B	582/672 (87%)	563 (97%)	19 (3%)	38	68
1	C	565/672 (84%)	547 (97%)	18 (3%)	39	69
1	D	597/672 (89%)	580 (97%)	17 (3%)	43	72
1	E	597/672 (89%)	568 (95%)	29 (5%)	25	59
1	F	589/672 (88%)	569 (97%)	20 (3%)	37	68
All	All	3528/4032 (88%)	3398 (96%)	130 (4%)	34	65

5 of 130 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	444	ARG
1	A	545	PHE
1	E	239	ARG
1	E	232	ARG
1	A	569	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	420	GLN
1	C	251	HIS
1	E	420	GLN
1	D	420	GLN
1	F	420	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 monosaccharides 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	689/775 (88%)	0.33	3 (0%) 92 90	32, 56, 100, 127	0
1	B	668/775 (86%)	0.56	39 (5%) 23 20	34, 76, 118, 132	0
1	C	648/775 (83%)	0.72	69 (10%) 6 7	39, 85, 150, 177	0
1	D	688/775 (88%)	0.67	65 (9%) 8 9	41, 84, 138, 160	0
1	E	688/775 (88%)	0.31	3 (0%) 92 90	41, 67, 111, 131	0
1	F	677/775 (87%)	0.75	68 (10%) 7 8	50, 95, 145, 169	0
All	All	4058/4650 (87%)	0.55	247 (6%) 21 19	32, 76, 135, 177	0

The worst 5 of 247 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	496	ILE	5.6
1	D	454	TRP	5.2
1	F	192	ASP	5.1
1	C	200	ILE	5.1
1	C	104	TYR	4.6

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.