



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:24 AM GMT

PDB ID : 3APZ  
Title : Apo form of Arabidopsis medium/long-chain length prenyl pyrophosphate synthase  
Authors : Hsieh, F.-L.; Chang, T.-H.; Ko, T.-P.; Wang, A.H.-J.  
Deposited on : 2010-10-24  
Resolution : 2.60 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

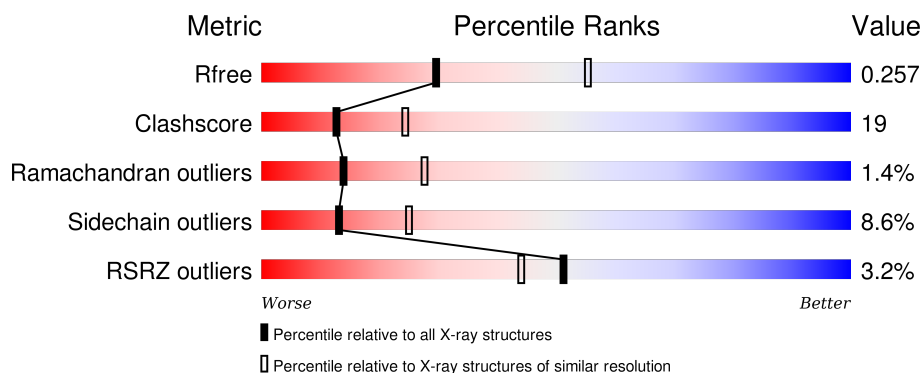
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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. 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	348	<div> <div>3%</div> <div>57%</div> <div>25%</div> <div>•</div> <div>14%</div> </div>
1	B	348	<div> <div>3%</div> <div>60%</div> <div>21%</div> <div>5%</div> <div>14%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4959 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 Geranyl diphosphate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	299	Total	C	N	O	S	0	0	0
			2293	1439	397	445	12			
1	B	299	Total	C	N	O	S	0	0	0
			2293	1439	397	445	12			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	EXPRESSION TAG	UNP Q9FT89
B	1	GLY	-	EXPRESSION TAG	UNP Q9FT89

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	175	Total	O	0	0
			175	175		
2	B	198	Total	O	0	0
			198	198		

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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:

3% 57% 25% 14%

GLY SER LEU VAL GLU GLU L8 D9 V14 A15 D16 E17 L20 L21 S22 N23 K24 L25 R26 E27 M28 V29 E32 V33 P34 LYS LEU ALA SER SER ALA GLU TYR PHE LYS ARG G47 V48 Q49 G50 K51 G52 F53 I57 L58 V59 T63 A64 L65 D66 V67 ARG TRP

PRO GLU ALA LEU ILE GLY GLU SER THR ASP ILE VAL T32 S33 E34 Q89 R90 G91 I92 I95 S103 L104 L105 V109 LEU ASP ALA ALA ASP THR ARG ARG GLY VAL GLY SER LEU ASN VAL M126 G127 N128 K129 M130 S131 V132 L133 A134 R141 L146 K150 N151

V154 L158 V162 V166 E172 I173 T174 S175 S176 T177 E178 Q179 R180 Y181 K189 I190 Y191 Y192 K193 S199 N200 S201 T208 T209 G210 Q211 T212 A213 E214 V215 A216 V217 F220 E221 R224 N225 L226 G227 L228 T242 T259 T260 A261 P262 T263 P271 G272

Chain B:

GLY	SER	LEU	VAL	GLU	GLU	GLU	LEU	LEU	D9	P10	F11	S12	L13	D16	L20	L21	S22	N23	K24	L25	R26	E27	M28	V29	L30	A31	P34	LYS	LEU	ALA	SER	ALA	ALA	GLU	TYR	PHE	PHE	LYS	ARG	G47	V48	T63	A64	L65	D66	V67	ARG	VAL	PRO	GLU	ALA	LEU	ILE	GLY	GLU	SER
THR	ILE	VAL	T82	S83	E84	L85	R86	V87	R90	G91	I95	D108	V109	L110	ASP	ALA	ASP	THR	ARG	GLY	VAL	GLY	SER	LEU	ASN	VAL	VAL	M126	G127	N128	K129	V132	R141	L146	T152	E153	V154	L158	A161	L165	T170	M171	E172	T173	T177											
V280	D288	L294	K298	Q301	A311	M312	I318	T324	S324	S322	I337	D338	L339	T340	H341	R342	R346	N347	K348	T277	Q272	L273	R274	V277	Q231	D234	T242	S245	L246	G247	K248	G249	S250	T259	T260	A261	P262	L263	L264	Q272	L273	R274	V277													

## 4 Data and refinement statistics

Property	Value	Source
Space group	F 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	148.51Å 150.13Å 176.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.60 23.83 – 2.60	Depositor EDS
% Data completeness (in resolution range)	93.1 (30.00-2.60) 93.1 (23.83-2.60)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.71 (at 2.60Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.206 , 0.256 0.206 , 0.257	Depositor DCC
$R_{free}$ test set	1371 reflections (4.86%)	DCC
Wilson B-factor (Å <sup>2</sup> )	39.0	Xtriage
Anisotropy	0.447	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 57.9	EDS
Estimated twinning fraction	0.020 for -k,-h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 28281 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4959	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.58% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.69	0/2316	0.80	1/3125 (0.0%)
1	B	0.69	0/2316	0.79	1/3125 (0.0%)
All	All	0.69	0/4632	0.80	2/6250 (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	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	158	LEU	CA-CB-CG	-7.19	98.77	115.30
1	B	158	LEU	CA-CB-CG	-5.59	102.43	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	191	TYR	Sidechain
1	B	191	TYR	Sidechain

## 5.2 Too-close contacts

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	2293	0	2345	101	0
1	B	2293	0	2345	81	0
2	A	175	0	0	2	0
2	B	198	0	0	6	0
All	All	4959	0	4690	179	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 179 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:259:ILE:HD11	1:A:277:VAL:HG11	1.42	0.98
1:B:211:GLN:HG2	1:B:215:VAL:HG11	1.49	0.93
1:B:190:THR:O	1:B:194:THR:HG22	1.69	0.93
1:A:176:SER:H	1:A:179:GLN:HE21	1.24	0.86
1:A:174:THR:HG22	1:A:174:THR:O	1.76	0.84

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	A	291/348 (84%)	280 (96%)	8 (3%)	3 (1%)	19	39
1	B	291/348 (84%)	275 (94%)	11 (4%)	5 (2%)	11	22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	582/696 (84%)	555 (95%)	19 (3%)	8 (1%)	14	28

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	128	ASN
1	B	128	ASN
1	A	29	VAL
1	B	127	GLY
1	B	273	LEU

### 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	A	249/289 (86%)	231 (93%)	18 (7%)	18	35
1	B	249/289 (86%)	224 (90%)	25 (10%)	9	18
All	All	498/578 (86%)	455 (91%)	43 (9%)	13	25

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	22	SER
1	B	90	ARG
1	B	324	THR
1	B	26	ARG
1	B	27	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	312	ASN
1	B	52	GLN
1	B	272	GLN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	301	GLN
1	B	211	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](#)

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	299/348 (85%)	-0.22	10 (3%) 50 43	23, 44, 93, 125	0
1	B	299/348 (85%)	-0.28	9 (3%) 54 47	20, 40, 98, 120	0
All	All	598/696 (85%)	-0.25	19 (3%) 51 44	20, 42, 96, 125	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	29	VAL	4.1
1	B	126	MET	4.1
1	A	47	GLY	3.5
1	A	34	PRO	3.2
1	A	82	THR	3.2

### 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.