



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 11:09 am BST

PDB ID : 4PPM  
Title : Crystal structure of PigE: a transaminase involved in the biosynthesis of 2-methyl-3-n-amyI-pyrrole (MAP) from *Serratia* sp. FS14  
Authors : Lou, X.D.; Ran, T.T.; Xu, D.Q.; Wang, W.W.  
Deposited on : 2014-02-27  
Resolution : 2.30 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

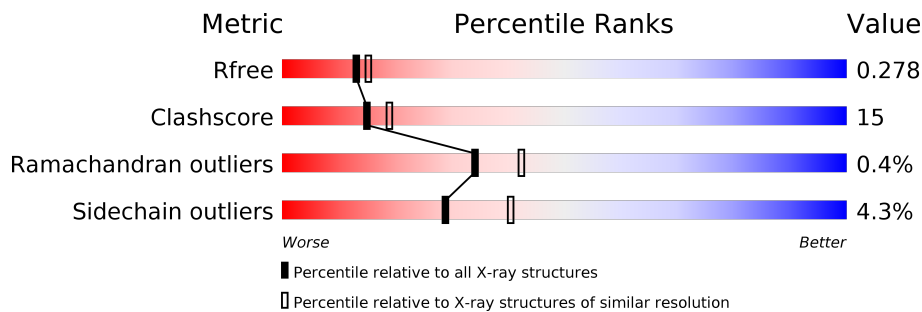
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	861	 38% 15% • 46%
1	B	861	 35% 17% • 46%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7493 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 Aminotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	464	3563	2265	617	657	24	0	0	0
1	B	462	3554	2263	615	652	24	0	0	0

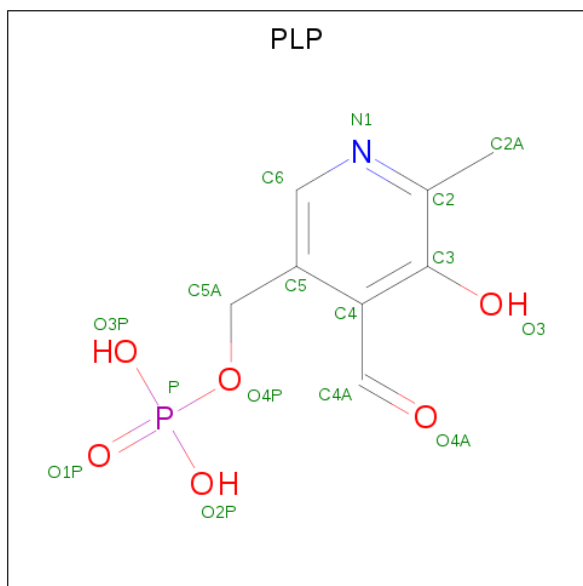
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	854	LEU	-	EXPRESSION TAG	UNP A0A059ZJX2
A	855	GLU	-	EXPRESSION TAG	UNP A0A059ZJX2
A	856	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
A	857	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
A	858	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
A	859	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
A	860	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
A	861	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
B	854	LEU	-	EXPRESSION TAG	UNP A0A059ZJX2
B	855	GLU	-	EXPRESSION TAG	UNP A0A059ZJX2
B	856	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
B	857	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
B	858	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
B	859	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
B	860	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2
B	861	HIS	-	EXPRESSION TAG	UNP A0A059ZJX2

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		

- Molecule 3 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: C<sub>8</sub>H<sub>10</sub>NO<sub>6</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	15	8	1	5	1	0	0
3	B	1	15	8	1	5	1	0	0

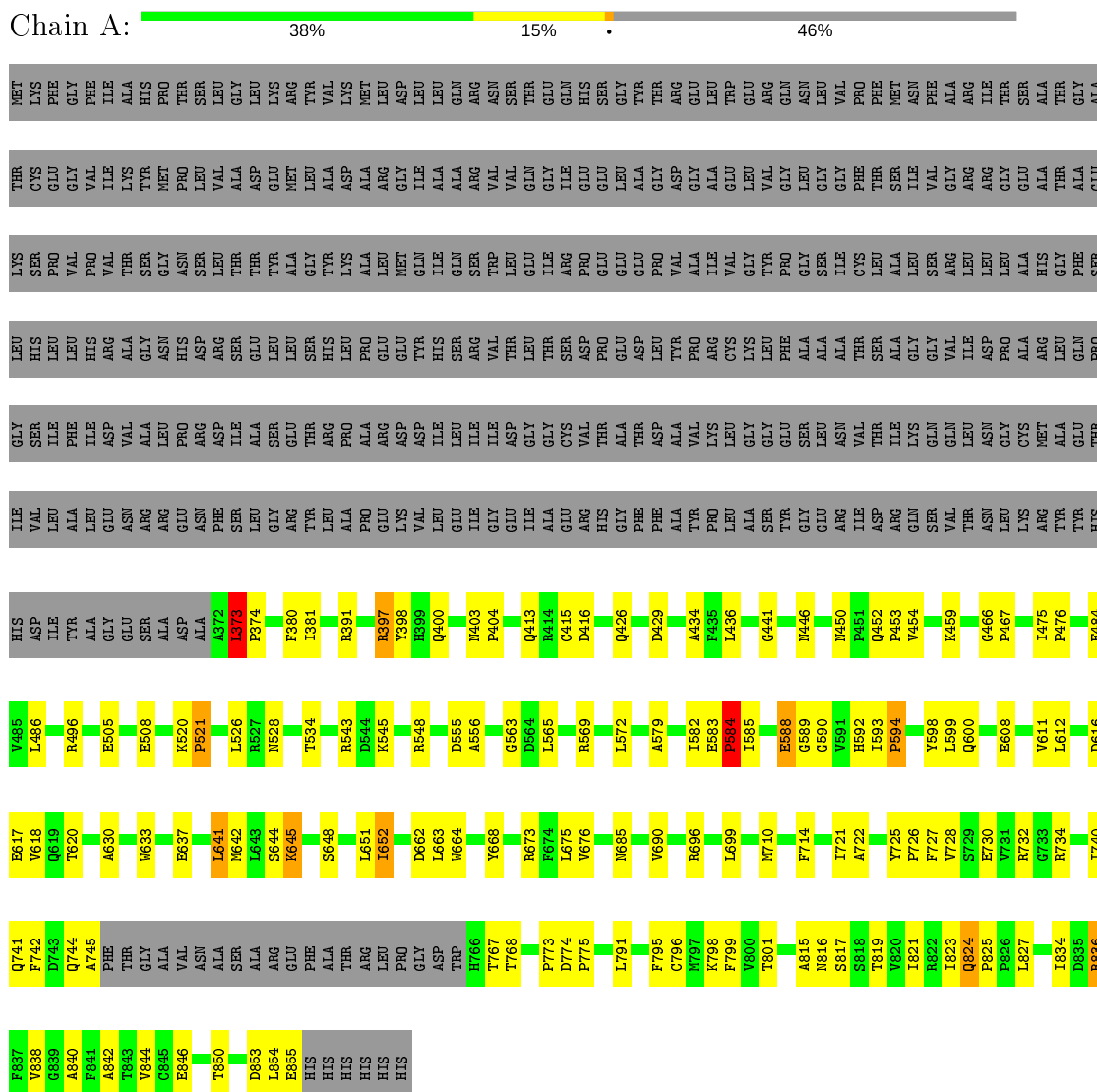
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	160	160	160	0	0
4	B	184	184	184	0	0

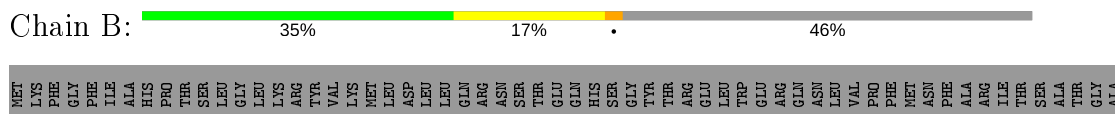
### 3 Residue-property plots

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. 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. 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: Aminotransferase



- Molecule 1: Aminotransferase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	228.10 Å 228.10 Å 67.10 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.89 – 2.30 19.89 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (19.89-2.30) 99.9 (19.89-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 2.30 Å)	Xtrriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.238 , 0.272 0.251 , 0.278	Depositor DCC
$R_{free}$ test set	4443 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.2	Xtrriage
Anisotropy	0.092	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 16.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.21$	Xtrriage
Estimated twinning fraction	0.340 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7493	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.01% 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: MG, PLP

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	1.41	7/3633 (0.2%)	0.94	19/4920 (0.4%)
1	B	1.36	6/3626 (0.2%)	0.95	20/4912 (0.4%)
All	All	1.39	13/7259 (0.2%)	0.95	39/9832 (0.4%)

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	505	GLU	CD-OE2	-5.65	1.19	1.25
1	A	508	GLU	CD-OE1	-5.60	1.19	1.25
1	A	508	GLU	CD-OE2	-5.57	1.19	1.25
1	B	404	PRO	N-CD	5.52	1.55	1.47
1	B	552	PRO	N-CD	5.44	1.55	1.47

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	702	ASP	CB-CG-OD1	7.09	124.68	118.30
1	A	590	GLY	N-CA-C	6.99	130.58	113.10
1	A	593	ILE	C-N-CD	6.29	141.61	128.40
1	B	725	TYR	C-N-CD	6.16	141.34	128.40
1	A	594	PRO	C-N-CD	6.02	141.04	128.40

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	3563	0	3567	100	0
1	B	3554	0	3558	132	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	15	0	6	0	0
3	B	15	0	6	0	0
4	A	160	0	0	5	0
4	B	184	0	0	6	0
All	All	7493	0	7137	216	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 15.

The worst 5 of 216 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:727:PHE:HB3	1:A:744:GLN:NE2	1.44	1.32
1:A:588:GLU:O	1:A:816:ASN:HB2	1.43	1.14
1:B:388:THR:HG22	4:B:1037:HOH:O	1.55	1.06
1:A:727:PHE:CB	1:A:744:GLN:HE21	1.69	1.05
1:A:727:PHE:HB3	1:A:744:GLN:HE21	0.89	1.04

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	460/861 (53%)	432 (94%)	26 (6%)	2 (0%)	34 42
1	B	458/861 (53%)	440 (96%)	16 (4%)	2 (0%)	34 42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	918/1722 (53%)	872 (95%)	42 (5%)	4 (0%)	34 42

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	645	LYS
1	B	645	LYS
1	A	644	SER
1	B	644	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	372/692 (54%)	360 (97%)	12 (3%)	39 54
1	B	371/692 (54%)	351 (95%)	20 (5%)	22 30
All	All	743/1384 (54%)	711 (96%)	32 (4%)	29 40

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	521	PRO
1	B	550	PHE
1	B	812	PHE
1	B	544	ASP
1	B	572	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	744	GLN
1	B	824	GLN
1	A	805	GLN
1	A	741	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	777	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](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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
3	PLP	B	902	1	15,15,16	2.68	8 (53%)	20,22,23	1.67	4 (20%)
3	PLP	A	902	1	15,15,16	2.86	9 (60%)	20,22,23	1.94	6 (30%)

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
3	PLP	B	902	1	-	1/6/6/8	0/1/1/1
3	PLP	A	902	1	-	2/6/6/8	0/1/1/1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	902	PLP	C5-C4	6.07	1.47	1.40
3	B	902	PLP	C5-C4	5.58	1.46	1.40
3	A	902	PLP	C3-C2	4.34	1.45	1.40
3	B	902	PLP	C3-C2	4.28	1.45	1.40
3	A	902	PLP	P-O3P	-3.78	1.40	1.54

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	902	PLP	C4A-C4-C5	5.05	126.13	120.94
3	B	902	PLP	O3P-P-O2P	3.12	119.54	107.64
3	B	902	PLP	O4P-C5A-C5	2.72	114.54	109.35
3	A	902	PLP	C3-C4-C5	-2.69	115.83	118.74
3	A	902	PLP	O3P-P-O2P	2.66	117.79	107.64

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	902	PLP	C6-C5-C5A-O4P
3	A	902	PLP	C4-C5-C5A-O4P
3	B	902	PLP	C5A-O4P-P-O1P

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

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.