



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 06:05 PM GMT

PDB ID : 1G78  
Title : X-RAY STRUCTURE OF ESCHERICHIA COLI PYRIDOXINE 5'-  
PHOSPHATE OXIDASE COMPLEXED WITH PYRIDOXAL 5'-  
PHOSPHATE AT 2.0 Å RESOLUTION  
Authors : Safo, M.K.; Musayev, F.N.; di Salvo, M.L.; Schirch, V.  
Deposited on : 2000-11-09  
Resolution : 2.20 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

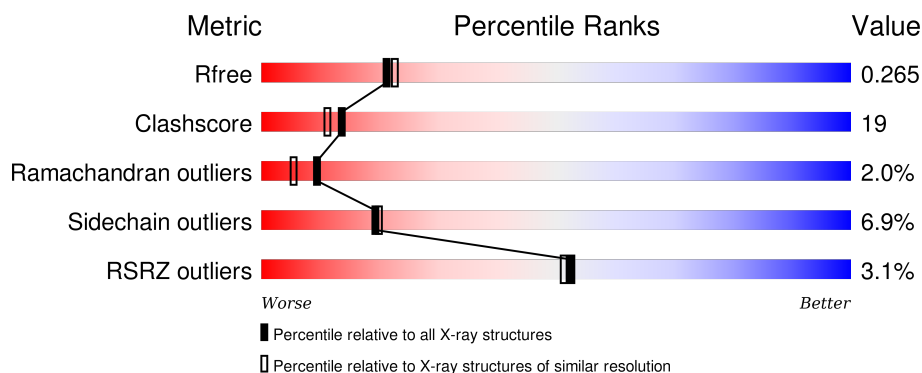
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	218	<div> <div>3%</div> <div>62%</div> <div>24%</div> <div>6%</div> <div>9%</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	PLP	A	270	-	-	-	X

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 1804 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.

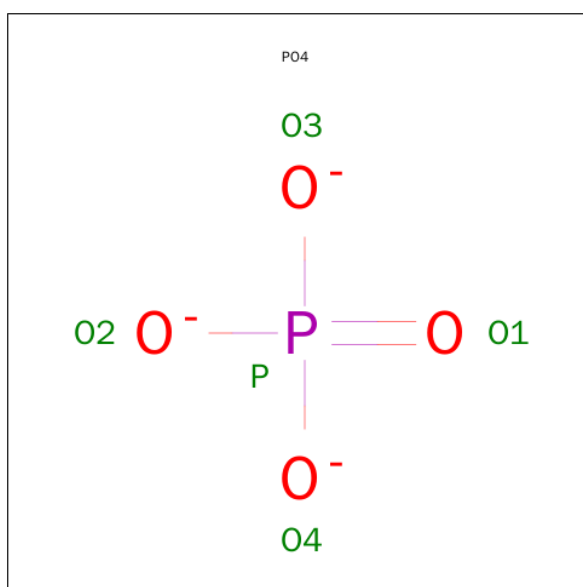
- Molecule 1 is a protein called PYRIDOXINE 5'-PHOSPHATE OXIDASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	199	1644	1048	301	289	1	5	0	1	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP P28225
A	53	MSE	MET	MODIFIED RESIDUE	UNP P28225
A	79	MSE	MET	MODIFIED RESIDUE	UNP P28225
A	113	MSE	MET	MODIFIED RESIDUE	UNP P28225
A	127	MSE	MET	MODIFIED RESIDUE	UNP P28225

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
2	A	1	5	4	1	0	0

*Continued on next page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	P	0	0
			5	4	1		

- 
- The image displays the chemical structure of Flavin Mononucleotide (FMN). It features an isoalloxazine ring system, which is a tricyclic aromatic heterocycle consisting of a benzene ring fused to a pyrimidine ring, which is further fused to a five-membered nitrogen-containing ring. The ring atoms are labeled: N1, N3, N5, N10, C2, C4, C6, C7, C8, C9, C10, C4A, C5A, C5B, C6A, C7A, C8A, C9A, C10A, C11, C12, C13, C14, C15, C16, C17, C18, C19, C20, C21, C22, C23, C24, C25, C26, C27, C28, C29, C30, C31, C32, C33, C34, C35, C36, C37, C38, C39, C40, C41, C42, C43, C44, C45, C46, C47, C48, C49, C50, C51, C52, C53, C54, C55, C56, C57, C58, C59, C60, C61, C62, C63, C64, C65, C66, C67, C68, C69, C70, C71, C72, C73, C74, C75, C76, C77, C78, C79, C80, C81, C82, C83, C84, C85, C86, C87, C88, C89, C90, C91, C92, C93, C94, C95, C96, C97, C98, C99, C100. The side chain is attached to the N10 position and consists of a ribityl chain (C1, C2, C3, C4) with a phosphate group (O1P, O2P, O3P, O4P) attached to the C4 position. The ribityl chain is shown in a 3D representation with wedged and dashed bonds indicating stereochemistry. The phosphate group is shown as a tetrahedral structure with a central phosphorus atom (P) bonded to four oxygen atoms (O1P, O2P, O3P, O4P). The overall structure is colored with green for the ring atoms, red for the side chain and phosphate group, and blue for the nitrogen atoms.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	17	4	9	1		

- PLP
- 
- The diagram shows the chemical structure of PLP (Pyridoxal phosphate). The central pyridine ring is colored green. The nitrogen atom is labeled N1 in blue. The carbon atoms are labeled C2, C3, C4, C5, and C6 in green. The aldehyde group is shown as C4A=O4A in green. The phosphate group is shown as C5A-O4P-P(O)(O1P)OH in red, with the phosphorus atom labeled P. The hydroxyl group is shown as C3-OH in red, with the oxygen atom labeled O3 in green. The labels are placed around the structure to identify specific atoms.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			16	8	1	6	1		
4	A	1	Total	C	N	O	P	0	0
			16	8	1	6	1		

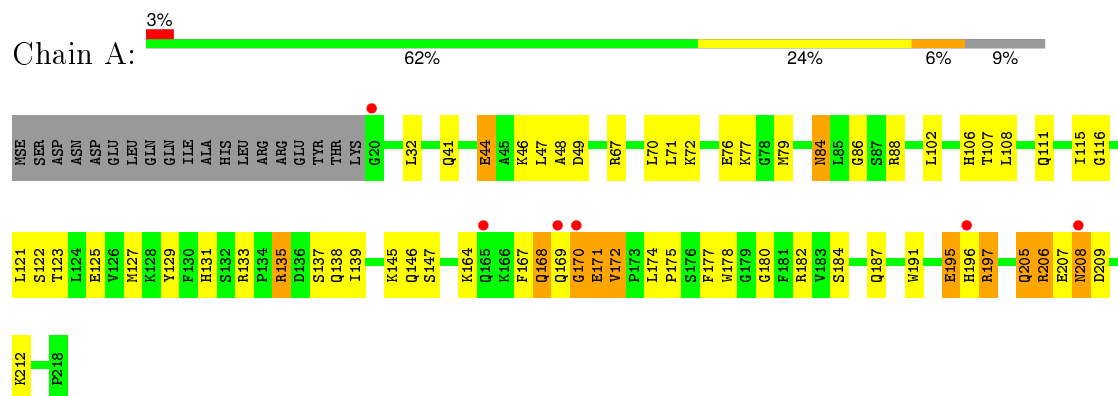
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	87	Total	O	0	0
			87	87		

### 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 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 ( $RSRZ > 2$ ). 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: PYRIDOXINE 5'-PHOSPHATE OXIDASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.40Å 63.40Å 124.44Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	55.00 – 2.20 50.23 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.4 (55.00-2.20) 98.9 (50.23-1.90)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	5.90	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.35 (at 1.90Å)	Xtriage
Refinement program	CNS 0.5	Depositor
R, $R_{free}$	0.218 , 0.265 0.218 , 0.265	Depositor DCC
$R_{free}$ test set	767 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	26.3	Xtriage
Anisotropy	0.379	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 48.3	EDS
Estimated twinning fraction	0.046 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 23258 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	1804	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.72% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: FMN, PO4, 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	0.52	0/1689	0.66	0/2271

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	A	1644	0	1630	63	0
2	A	10	0	0	0	0
3	A	31	0	19	1	0
4	A	32	0	16	4	0
5	A	87	0	0	4	0
All	All	1804	0	1665	63	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 63 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:84:ASN:HD22	1:A:86:GLY:H	1.23	0.84
1:A:167:PHE:C	1:A:169:GLN:H	1.77	0.84
1:A:123:THR:HG22	5:A:327:HOH:O	1.79	0.82
1:A:135:ARG:CD	1:A:135:ARG:H	1.95	0.77
1:A:84:ASN:ND2	1:A:86:GLY:H	1.83	0.76

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	198/218 (91%)	184 (93%)	10 (5%)	4 (2%)	<b>9</b> <b>5</b>

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	171	GLU
1	A	172	VAL
1	A	208	ASN
1	A	170	GLY

### 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	175/187 (94%)	163 (93%)	12 (7%)	<b>19</b> <b>20</b>

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	137	SER
1	A	168	GLN
1	A	197	ARG
1	A	135	ARG
1	A	195	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	106	HIS
1	A	205	GLN
1	A	131	HIS
1	A	84	ASN
1	A	196	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

5 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
3	FMN	A	250	-	32,33,33	2.69	14 (43%)	34,50,50	3.22	11 (32%)
4	PLP	A	260	-	16,16,16	2.63	6 (37%)	21,23,23	1.21	2 (9%)
4	PLP	A	270	-	16,16,16	2.37	7 (43%)	21,23,23	1.21	2 (9%)
2	PO4	A	280	-	4,4,4	1.34	0	6,6,6	0.24	0
2	PO4	A	290	-	4,4,4	0.39	0	6,6,6	0.24	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
3	FMN	A	250	-	-	0/18/18/18	0/3/3/3
4	PLP	A	260	-	-	0/8/8/8	0/1/1/1
4	PLP	A	270	-	-	0/8/8/8	0/1/1/1
2	PO4	A	280	-	-	0/0/0/0	0/0/0/0
2	PO4	A	290	-	-	0/0/0/0	0/0/0/0

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	250	FMN	P-O2P	-4.52	1.39	1.54
3	A	250	FMN	C7M-C7	2.13	1.55	1.51
4	A	270	PLP	P-O3P	2.24	1.62	1.54
3	A	250	FMN	C2-N3	2.24	1.42	1.38
4	A	270	PLP	C3-C2	2.35	1.42	1.40

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	250	FMN	C4-C4A-C10	-5.78	116.24	119.94
3	A	250	FMN	C4A-C10-N10	-5.74	116.35	120.52
3	A	250	FMN	N3-C2-N1	-5.20	118.93	127.69
3	A	250	FMN	C4A-C4-N3	-4.32	117.88	123.52
4	A	270	PLP	C5-C6-N1	-2.55	119.41	123.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	250	FMN	1	0
4	A	260	PLP	2	0
4	A	270	PLP	2	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	195/218 (89%)	-0.32	6 (3%) 52 51	21, 32, 54, 62	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	20	GLY	4.8
1	A	169	GLN	3.7
1	A	170	GLY	3.3
1	A	208	ASN	3.2
1	A	196	HIS	2.6

### 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 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	PLP	A	270	16/16	0.79	0.17	3.37	52,53,60,60	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	PO4	A	280	5/5	0.99	0.12	1.43	20,22,24,24	0
3	FMN	A	250	31/31	0.97	0.12	0.93	24,29,31,32	0
4	PLP	A	260	16/16	0.88	0.13	0.60	47,49,56,58	0
2	PO4	A	290	5/5	0.93	0.18	-	65,65,66,66	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.