



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 07:53 pm BST

PDB ID : 1K23  
Title : Inorganic Pyrophosphatase (Family II) from *Bacillus subtilis*  
Authors : Ahn, S.; Milner, A.J.; Futterer, K.; Konopka, M.; Ilias, M.; Young, T.W.;  
White, S.A.  
Deposited on : 2001-09-26  
Resolution : 3.00 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)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

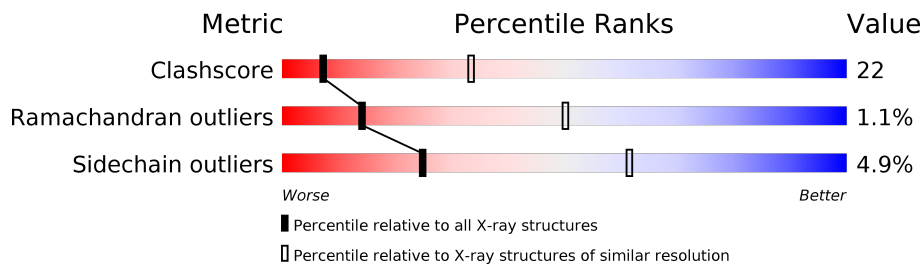
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	309	
1	B	309	
1	C	309	
1	D	309	

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 8973 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 Manganese-dependent inorganic pyrophosphatase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	307	2371	1496	392	475	3	5	0	0	0
1	B	307	2371	1496	392	475	3	5	0	0	0
1	C	308	2376	1499	393	476	3	5	0	0	0
1	D	231	1789	1123	296	362	3	5	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP P37487
A	126	MSE	MET	MODIFIED RESIDUE	UNP P37487
A	142	MSE	MET	MODIFIED RESIDUE	UNP P37487
A	184	MSE	MET	MODIFIED RESIDUE	UNP P37487
A	307	MSE	MET	MODIFIED RESIDUE	UNP P37487
B	1	MSE	MET	MODIFIED RESIDUE	UNP P37487
B	126	MSE	MET	MODIFIED RESIDUE	UNP P37487
B	142	MSE	MET	MODIFIED RESIDUE	UNP P37487
B	184	MSE	MET	MODIFIED RESIDUE	UNP P37487
B	307	MSE	MET	MODIFIED RESIDUE	UNP P37487
C	1	MSE	MET	MODIFIED RESIDUE	UNP P37487
C	126	MSE	MET	MODIFIED RESIDUE	UNP P37487
C	142	MSE	MET	MODIFIED RESIDUE	UNP P37487
C	184	MSE	MET	MODIFIED RESIDUE	UNP P37487
C	307	MSE	MET	MODIFIED RESIDUE	UNP P37487
D	1	MSE	MET	MODIFIED RESIDUE	UNP P37487
D	126	MSE	MET	MODIFIED RESIDUE	UNP P37487
D	142	MSE	MET	MODIFIED RESIDUE	UNP P37487
D	184	MSE	MET	MODIFIED RESIDUE	UNP P37487
D	307	MSE	MET	MODIFIED RESIDUE	UNP P37487

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total 2	Mn 2	0	0
2	A	2	Total 2	Mn 2	0	0
2	D	2	Total 2	Mn 2	0	0
2	C	2	Total 2	Mn 2	0	0

- Molecule 3 is water.

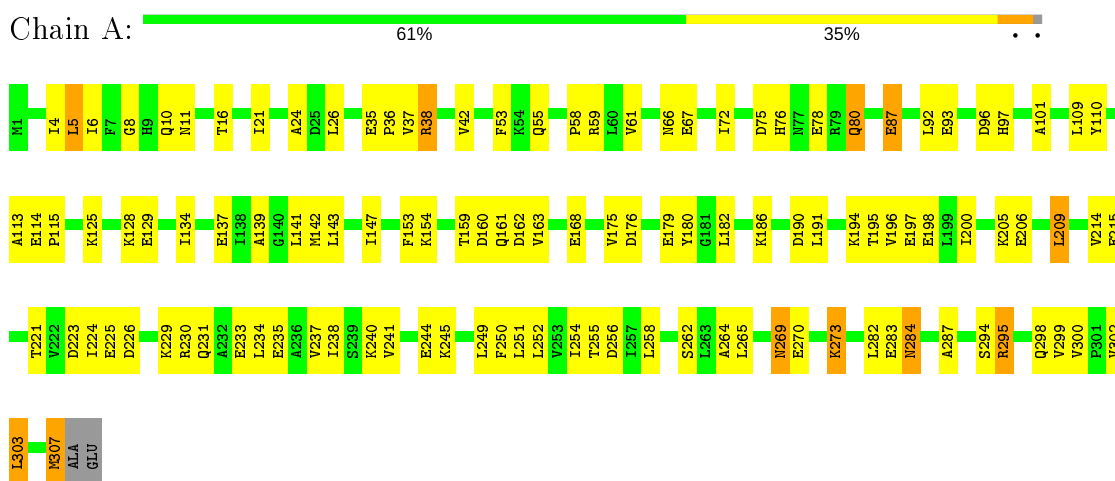
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	24	Total 24	O 24	0	0
3	B	9	Total 9	O 9	0	0
3	C	13	Total 13	O 13	0	0
3	D	12	Total 12	O 12	0	0

### 3 Residue-property plots

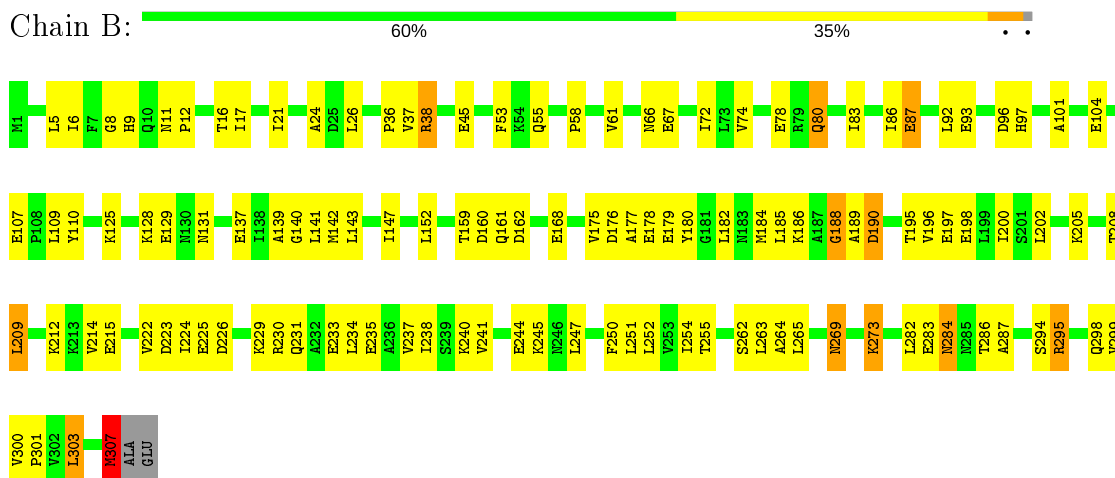
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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.

Note EDS was not executed.

- Molecule 1: Manganese-dependent inorganic pyrophosphatase



- Molecule 1: Manganese-dependent inorganic pyrophosphatase



- Molecule 1: Manganese-dependent inorganic pyrophosphatase





## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.29Å 130.56Å 150.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 3.00	Depositor
% Data completeness (in resolution range)	97.9 (19.96-3.00)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.04	Depositor
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.260 , 0.293	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	8973	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

## 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: MN

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	0.50	1/2392 (0.0%)	0.62	0/3230
1	B	0.47	1/2392 (0.0%)	0.63	0/3230
1	C	0.45	0/2397	0.61	0/3237
1	D	0.57	1/1804 (0.1%)	0.65	0/2435
All	All	0.49	3/8985 (0.0%)	0.63	0/12132

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	307	MSE	CG-SE	8.75	2.25	1.95
1	A	307	MSE	CG-SE	7.12	2.19	1.95
1	B	307	MSE	CG-SE	6.30	2.16	1.95

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	2371	0	2420	102	0
1	B	2371	0	2420	129	0
1	C	2376	0	2425	87	0
1	D	1789	0	1789	92	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	24	0	0	4	0
3	B	9	0	0	9	0
3	C	13	0	0	2	0
3	D	12	0	0	3	0
All	All	8973	0	9054	404	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 22.

The worst 5 of 404 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:307:MSE:CE	1:A:307:MSE:SE	2.15	1.45
1:B:307:MSE:SE	1:B:307:MSE:CG	2.16	1.43
1:B:307:MSE:SE	1:B:307:MSE:CE	2.15	1.43
1:A:307:MSE:CG	1:A:307:MSE:SE	2.19	1.41
1:D:307:MSE:SE	1:D:307:MSE:CE	2.23	1.37

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	305/309 (99%)	284 (93%)	19 (6%)	2 (1%)	22 60
1	B	305/309 (99%)	281 (92%)	19 (6%)	5 (2%)	9 40
1	C	306/309 (99%)	283 (92%)	21 (7%)	2 (1%)	22 60

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	223/309 (72%)	197 (88%)	22 (10%)	4 (2%)	8	37
All	All	1139/1236 (92%)	1045 (92%)	81 (7%)	13 (1%)	14	50

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	202	LEU
1	B	190	ASP
1	D	192	SER
1	D	226	ASP
1	D	254	ILE

### 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	261/257 (102%)	249 (95%)	12 (5%)	27	64
1	B	261/257 (102%)	247 (95%)	14 (5%)	22	57
1	C	261/257 (102%)	249 (95%)	12 (5%)	27	64
1	D	198/257 (77%)	188 (95%)	10 (5%)	24	60
All	All	981/1028 (95%)	933 (95%)	48 (5%)	25	61

5 of 48 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	269	ASN
1	C	5	LEU
1	D	190	ASP
1	B	273	LYS
1	B	303	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	218	GLN
1	B	269	ASN
1	D	33	ASN
1	B	66	ASN
1	B	69	ASN

### 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 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.