



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 6, 2024 – 10:28 AM EST

PDB ID : 2FOR  
Title : Crystal Structure of the Shigella flexneri Farnesyl Pyrophosphate Synthase Complex with an Isopentenyl Pyrophosphate  
Authors : Minasov, G.; Brunzelle, J.S.; Shuvalova, L.; Collart, F.R.; Joachimiak, A.; Anderson, W.F.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2006-01-13  
Resolution : 2.00 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 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
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

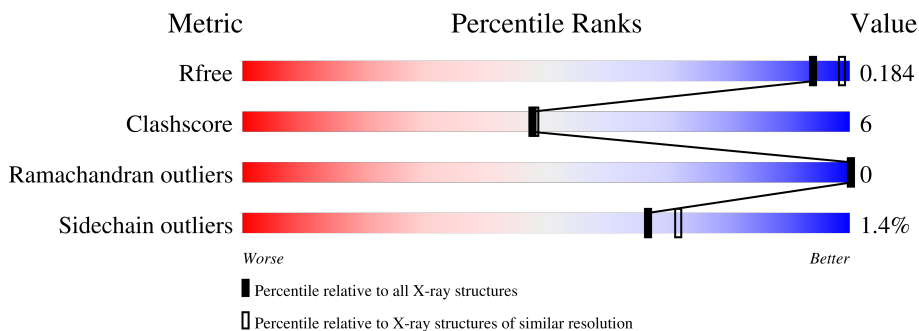
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	323	 73%      14%      • 12%
1	B	323	 76%      10%      • 13%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4743 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 Geranyltranstransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	283	2156	1346	385	414	11	0	3	0
1	B	281	2170	1355	386	416	13	0	6	0

There are 48 discrepancies between the modelled and reference sequences:

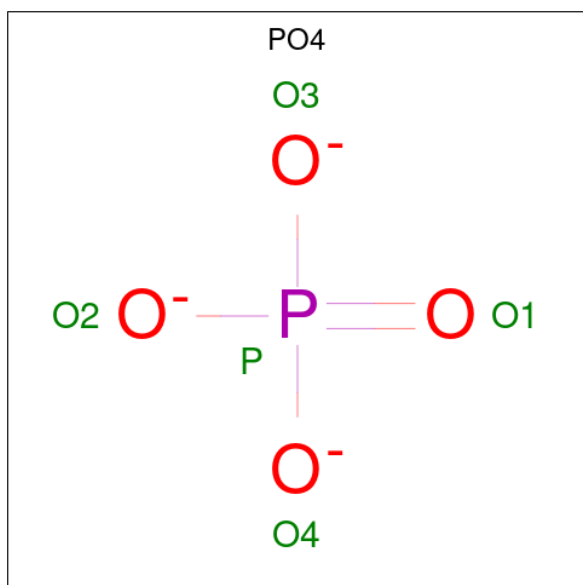
Chain	Residue	Modelled	Actual	Comment	Reference
A	-23	MET	-	expression tag	GB 24050587
A	-22	HIS	-	expression tag	GB 24050587
A	-21	HIS	-	expression tag	GB 24050587
A	-20	HIS	-	expression tag	GB 24050587
A	-19	HIS	-	expression tag	GB 24050587
A	-18	HIS	-	expression tag	GB 24050587
A	-17	HIS	-	expression tag	GB 24050587
A	-16	SER	-	expression tag	GB 24050587
A	-15	SER	-	expression tag	GB 24050587
A	-14	GLY	-	expression tag	GB 24050587
A	-13	VAL	-	expression tag	GB 24050587
A	-12	ASP	-	expression tag	GB 24050587
A	-11	LEU	-	expression tag	GB 24050587
A	-10	GLY	-	expression tag	GB 24050587
A	-9	THR	-	expression tag	GB 24050587
A	-8	GLU	-	expression tag	GB 24050587
A	-7	ASN	-	expression tag	GB 24050587
A	-6	LEU	-	expression tag	GB 24050587
A	-5	TYR	-	expression tag	GB 24050587
A	-4	PHE	-	expression tag	GB 24050587
A	-3	GLN	-	expression tag	GB 24050587
A	-2	SER	-	expression tag	GB 24050587
A	-1	ASN	-	expression tag	GB 24050587
A	0	ALA	-	expression tag	GB 24050587
B	-23	MET	-	expression tag	GB 24050587

*Continued on next page...*

Continued from previous page...

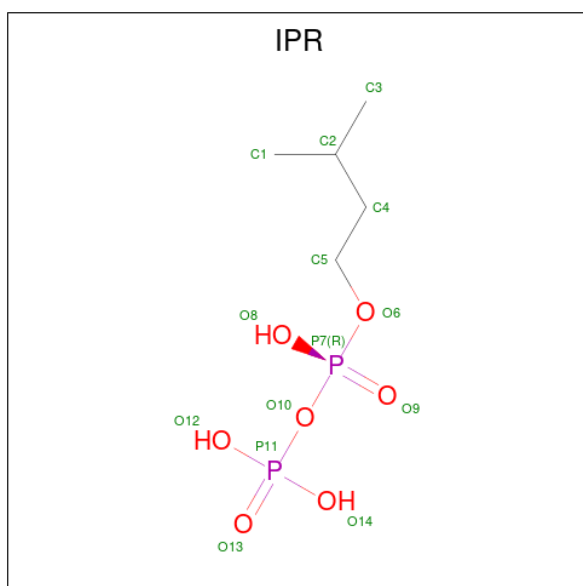
Chain	Residue	Modelled	Actual	Comment	Reference
B	-22	HIS	-	expression tag	GB 24050587
B	-21	HIS	-	expression tag	GB 24050587
B	-20	HIS	-	expression tag	GB 24050587
B	-19	HIS	-	expression tag	GB 24050587
B	-18	HIS	-	expression tag	GB 24050587
B	-17	HIS	-	expression tag	GB 24050587
B	-16	SER	-	expression tag	GB 24050587
B	-15	SER	-	expression tag	GB 24050587
B	-14	GLY	-	expression tag	GB 24050587
B	-13	VAL	-	expression tag	GB 24050587
B	-12	ASP	-	expression tag	GB 24050587
B	-11	LEU	-	expression tag	GB 24050587
B	-10	GLY	-	expression tag	GB 24050587
B	-9	THR	-	expression tag	GB 24050587
B	-8	GLU	-	expression tag	GB 24050587
B	-7	ASN	-	expression tag	GB 24050587
B	-6	LEU	-	expression tag	GB 24050587
B	-5	TYR	-	expression tag	GB 24050587
B	-4	PHE	-	expression tag	GB 24050587
B	-3	GLN	-	expression tag	GB 24050587
B	-2	SER	-	expression tag	GB 24050587
B	-1	ASN	-	expression tag	GB 24050587
B	0	ALA	-	expression tag	GB 24050587

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



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

- Molecule 3 is ISOPENTYL PYROPHOSPHATE (three-letter code: IPR) (formula:  $C_5H_{14}O_7P_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O P 14 5 7 2	0	1
3	B	1	Total C O P 14 5 7 2	0	1

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	161	Total O 163 163	0	7

*Continued on next page...*

*Continued from previous page...*

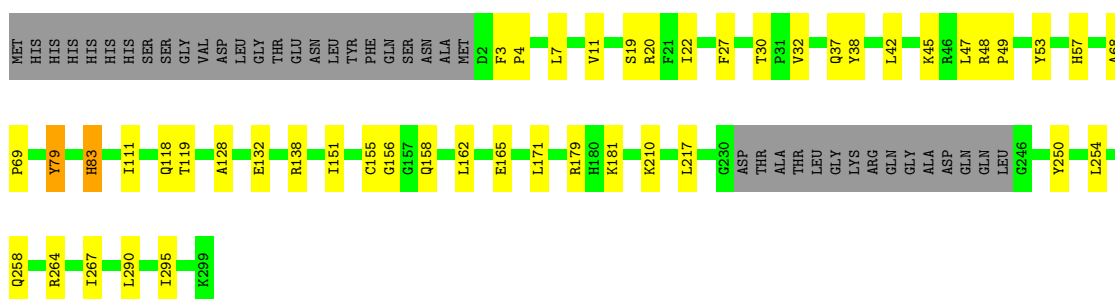
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	B	187	Total 191	O 191	0	12

### 3 Residue-property plots [i](#)


These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

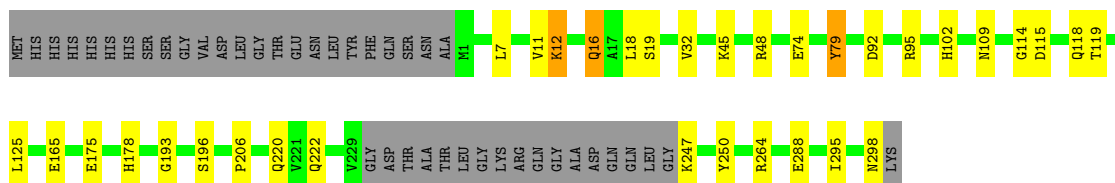
- Molecule 1: Geranyltranstransferase

Chain A:  73% 14% 12%



- Molecule 1: Geranyltranstransferase

Chain B:  76% 10% 13%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.66Å 75.66Å 216.66Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	65.51 – 2.00 24.77 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.7 (65.51-2.00) 97.7 (24.77-2.00)	Depositor EDS
$R_{merge}$	0.02	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.60 (at 1.99Å)	Xtrriage
Refinement program	REFMAC 5.2	Depositor
R, $R_{free}$	0.181 , 0.233 0.186 , 0.184	Depositor DCC
$R_{free}$ test set	2301 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.5	Xtrriage
Anisotropy	0.164	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 50.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	0.079 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4743	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.81% 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: IPR, PO4

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	1/2187 (0.0%)	0.67	0/2958
1	B	0.69	9/2201 (0.4%)	0.72	2/2978 (0.1%)
All	All	0.61	10/4388 (0.2%)	0.69	2/5936 (0.0%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	12	LYS	C-O	8.56	1.39	1.23
1	B	16[A]	GLN	C-O	7.91	1.38	1.23
1	B	16[B]	GLN	C-O	7.91	1.38	1.23
1	A	155	CYS	CB-SG	-7.03	1.70	1.82
1	B	16[A]	GLN	CD-OE1	6.06	1.37	1.24

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	16[A]	GLN	O-C-N	5.41	131.35	122.70
1	B	16[B]	GLN	O-C-N	5.41	131.35	122.70

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	2156	0	2162	31	0
1	B	2170	0	2176	25	0
2	A	25	0	0	1	0
2	B	10	0	0	0	0
3	A	14	0	10	1	0
3	B	14	0	10	4	0
4	A	163	0	0	4	0
4	B	191	0	0	2	0
All	All	4743	0	4358	53	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 6.

The worst 5 of 53 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:162:LEU:O	1:A:165:GLU:HG2	1.85	0.77
1:A:27:PHE:O	1:A:30:THR:HG22	1.88	0.73
1:A:132:GLU:HB2	4:A:413:HOH:O	1.91	0.70
1:A:83:HIS:HD2	1:B:115:ASP:OD2	1.75	0.70
1:B:165:GLU:HG2	1:B:247:LYS:HE2	1.77	0.67

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	282/323 (87%)	276 (98%)	6 (2%)	0	100	100
1	B	283/323 (88%)	280 (99%)	3 (1%)	0	100	100
All	All	565/646 (88%)	556 (98%)	9 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	224/254 (88%)	219 (98%)	5 (2%)	52	55
1	B	227/254 (89%)	226 (100%)	1 (0%)	91	93
All	All	451/508 (89%)	445 (99%)	6 (1%)	67	74

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	210	LYS
1	A	290	LEU
1	B	79	TYR
1	A	79	TYR
1	A	20	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	118	GLN
1	B	109	ASN
1	A	296	GLN
1	B	102	HIS
1	A	178	HIS

### 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](#)

9 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	IPR	A	305[B]	-	11,13,13	1.18	1 (9%)	16,19,19	1.26	1 (6%)
2	PO4	B	300	-	4,4,4	0.84	0	6,6,6	0.66	0
2	PO4	A	300[A]	-	4,4,4	1.07	0	6,6,6	0.56	0
3	IPR	B	302[B]	-	11,13,13	1.19	1 (9%)	16,19,19	1.28	1 (6%)
2	PO4	A	302[A]	-	4,4,4	0.88	0	6,6,6	0.57	0
2	PO4	A	304	-	4,4,4	0.83	0	6,6,6	0.50	0
2	PO4	A	301[A]	-	4,4,4	0.82	0	6,6,6	0.38	0
2	PO4	B	301[A]	-	4,4,4	0.84	0	6,6,6	0.46	0
2	PO4	A	303	-	4,4,4	0.86	0	6,6,6	0.52	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	IPR	A	305[B]	-	-	3/13/13/13	-
3	IPR	B	302[B]	-	-	6/13/13/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	305[B]	IPR	C1-C2	-3.47	1.32	1.51
3	B	302[B]	IPR	C1-C2	-3.46	1.32	1.51

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	302[B]	IPR	P7-O10-P11	-3.05	122.36	132.83
3	A	305[B]	IPR	P7-O10-P11	-3.02	122.46	132.83

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	305[B]	IPR	C2-C4-C5-O6
3	B	302[B]	IPR	P7-O10-P11-O12
3	B	302[B]	IPR	C2-C4-C5-O6
3	B	302[B]	IPR	C5-O6-P7-O8
3	B	302[B]	IPR	C1-C2-C4-C5

There are no ring outliers.

3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	305[B]	IPR	1	0
3	B	302[B]	IPR	4	0
2	A	303	PO4	1	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

### 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.