



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2023 – 09:16 AM EDT

PDB ID : 7TYE  
Title : The crystal structure of 3,4-dihydroxy-2-butanone 4-phosphate synthase mutant (G108S) from E. Coli  
Authors : Tan, K.; Perkovich, P.; Joachimiak, A.  
Deposited on : 2022-02-12  
Resolution : 1.94 Å (reported)

This is a Full wwPDB X-ray Structure Validation 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  
Xtrriage (Phenix) : 1.13  
EDS : **FAILED**  
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) : 2.35.1

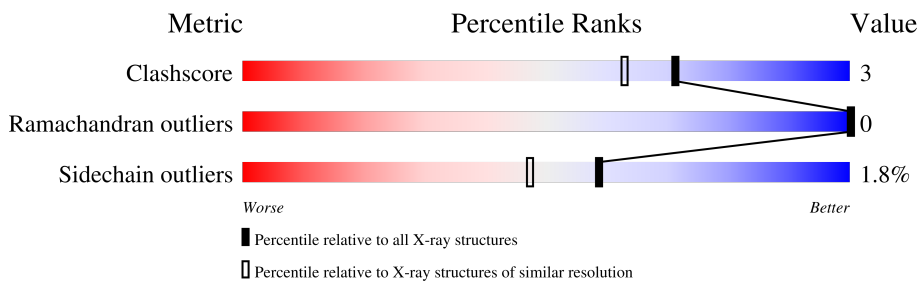
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.94 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	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)

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  $\leq 5\%$

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	226	
1	B	226	
1	C	226	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 4870 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 3,4-dihydroxy-2-butanone 4-phosphate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	213	1599	983	285	318	13	0	0	0
1	B	214	1604	986	286	319	13	0	0	0
1	C	213	1599	983	285	318	13	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	108	SER	GLY	engineered mutation	UNP C3SU87
A	218	GLY	-	expression tag	UNP C3SU87
A	219	SER	-	expression tag	UNP C3SU87
A	220	SER	-	expression tag	UNP C3SU87
A	221	HIS	-	expression tag	UNP C3SU87
A	222	HIS	-	expression tag	UNP C3SU87
A	223	HIS	-	expression tag	UNP C3SU87
A	224	HIS	-	expression tag	UNP C3SU87
A	225	HIS	-	expression tag	UNP C3SU87
A	226	HIS	-	expression tag	UNP C3SU87
B	108	SER	GLY	engineered mutation	UNP C3SU87
B	218	GLY	-	expression tag	UNP C3SU87
B	219	SER	-	expression tag	UNP C3SU87
B	220	SER	-	expression tag	UNP C3SU87
B	221	HIS	-	expression tag	UNP C3SU87
B	222	HIS	-	expression tag	UNP C3SU87
B	223	HIS	-	expression tag	UNP C3SU87
B	224	HIS	-	expression tag	UNP C3SU87
B	225	HIS	-	expression tag	UNP C3SU87
B	226	HIS	-	expression tag	UNP C3SU87
C	108	SER	GLY	engineered mutation	UNP C3SU87
C	218	GLY	-	expression tag	UNP C3SU87
C	219	SER	-	expression tag	UNP C3SU87

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	220	SER	-	expression tag	UNP C3SU87
C	221	HIS	-	expression tag	UNP C3SU87
C	222	HIS	-	expression tag	UNP C3SU87
C	223	HIS	-	expression tag	UNP C3SU87
C	224	HIS	-	expression tag	UNP C3SU87
C	225	HIS	-	expression tag	UNP C3SU87
C	226	HIS	-	expression tag	UNP C3SU87

- Molecule 2 is water.


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	33	Total O 33 33	0	0
2	B	20	Total O 20 20	0	0
2	C	15	Total O 15 15	0	0

### 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.


Note EDS failed to run properly.

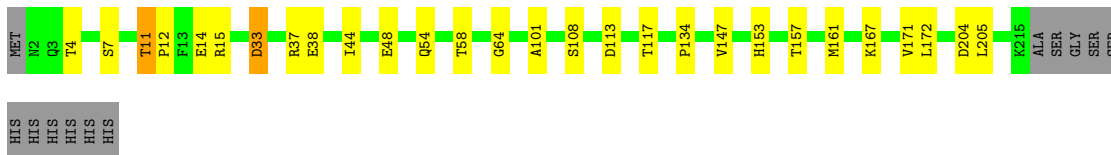
- Molecule 1: 3,4-dihydroxy-2-butanone 4-phosphate synthase

Chain A:  86% 7% • 6%




- Molecule 1: 3,4-dihydroxy-2-butanone 4-phosphate synthase

Chain B:  82% 12% • 5%



- Molecule 1: 3,4-dihydroxy-2-butanone 4-phosphate synthase

Chain C:  85% 8% 6%



## 4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	164.99Å 108.51Å 47.44Å 90.00° 98.74° 90.00°	Depositor
Resolution (Å)	45.20 – 1.94	Depositor
% Data completeness (in resolution range)	98.2 (45.20-1.94)	Depositor
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.75 (at 1.94Å)	Xtrriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, $R_{free}$	0.196 , 0.224	Depositor
Wilson B-factor (Å <sup>2</sup> )	46.6	Xtrriage
Anisotropy	0.191	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	4870	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	66.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% 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](#)

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	0.28	0/1620	0.49	1/2195 (0.0%)
1	B	0.26	0/1625	0.45	0/2202
1	C	0.27	0/1620	0.47	0/2195
All	All	0.27	0/4865	0.47	1/6592 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	172	LEU	CA-CB-CG	5.23	127.33	115.30

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	1599	0	1573	8	0
1	B	1604	0	1575	15	0
1	C	1599	0	1573	12	0
2	A	33	0	0	0	0
2	B	20	0	0	0	0
2	C	15	0	0	0	0
All	All	4870	0	4721	33	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 3.

All (33) 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:B:11:THR:HG22	1:B:12:PRO:HD2	1.81	0.62
1:B:15:ARG:NH1	1:B:204:ASP:OD2	2.39	0.56
1:C:100:GLU:OE2	1:C:108:SER:OG	2.21	0.56
1:B:11:THR:OG1	1:B:14:GLU:OE2	2.20	0.53
1:A:34:ASP:OD2	1:A:39:ASN:ND2	2.41	0.52
1:B:134:PRO:HD3	1:C:84:VAL:HG21	1.91	0.52
1:A:48:GLU:OE2	1:A:167:LYS:N	2.39	0.51
1:C:157:THR:O	1:C:161:MET:HG2	2.12	0.50
1:C:54:GLN:O	1:C:58:THR:HG23	2.12	0.50
1:C:30:MET:HB2	1:C:199:LEU:HD22	1.94	0.50
1:B:48:GLU:OE2	1:B:167:LYS:N	2.35	0.50
1:B:108:SER:HA	1:B:113:ASP:HB3	1.93	0.50
1:B:153:HIS:HB3	1:B:172:LEU:HD11	1.94	0.49
1:B:54:GLN:O	1:B:58:THR:HG23	2.13	0.49
1:A:38:GLU:OE1	1:A:184:ARG:NE	2.44	0.48
1:A:189:ILE:HA	1:A:199:LEU:HD11	1.94	0.48
1:B:64:GLY:HA3	1:C:64:GLY:HA3	1.96	0.47
1:C:48:GLU:OE2	1:C:167:LYS:N	2.44	0.47
1:C:34:ASP:OD1	1:C:39:ASN:ND2	2.47	0.46
1:C:108:SER:HA	1:C:113:ASP:HB3	1.97	0.46
1:B:33:ASP:OD2	1:B:33:ASP:N	2.49	0.46
1:A:30:MET:HG3	1:A:192:ALA:HB2	1.99	0.45
1:B:147:VAL:HG21	1:B:205:LEU:HG	1.99	0.44
1:B:4:THR:HG22	1:B:7:SER:OG	2.18	0.43
1:B:101:ALA:HB2	1:B:117:THR:HA	1.99	0.43
1:C:30:MET:HG3	1:C:192:ALA:HB2	2.01	0.43
1:B:44:ILE:HA	1:B:171:VAL:O	2.18	0.43
1:A:20:LEU:O	1:A:24:ARG:HG3	2.20	0.42
1:C:32:LEU:HG	1:C:199:LEU:HD11	2.01	0.42
1:B:157:THR:O	1:B:161:MET:HG2	2.20	0.41
1:C:101:ALA:HB2	1:C:117:THR:HA	2.01	0.41
1:A:101:ALA:HB2	1:A:117:THR:HA	2.03	0.40
1:A:108:SER:HA	1:A:113:ASP:HB3	2.03	0.40

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	211/226 (93%)	208 (99%)	3 (1%)	0	100	100
1	B	212/226 (94%)	206 (97%)	6 (3%)	0	100	100
1	C	211/226 (93%)	207 (98%)	4 (2%)	0	100	100
All	All	634/678 (94%)	621 (98%)	13 (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	169/181 (93%)	166 (98%)	3 (2%)	59	47
1	B	169/181 (93%)	165 (98%)	4 (2%)	49	36
1	C	169/181 (93%)	167 (99%)	2 (1%)	71	64
All	All	507/543 (93%)	498 (98%)	9 (2%)	59	47

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	85	GLU
1	A	172	LEU
1	A	199	LEU
1	B	11	THR
1	B	33	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	37	ARG
1	B	38	GLU
1	C	184	ARG
1	C	199	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

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

### 6.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

### 6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

### 6.4 Ligands

EDS failed to run properly - this section is therefore empty.

### 6.5 Other polymers

EDS failed to run properly - this section is therefore empty.