



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 30, 2024 – 10:14 AM JST

PDB ID : 8JG3  
Title : Biosynthetic thiolase from *Clostridium kluyveri*  
Authors : Kim, E.-J.; Seo, H.; Kim, K.-J.  
Deposited on : 2023-05-19  
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  
Xtriage (Phenix) : 1.13  
EDS : 2.36.2  
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) : 2.36.2

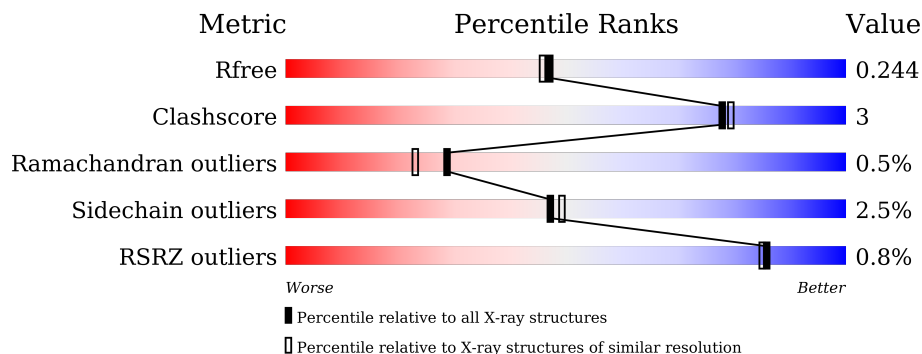
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

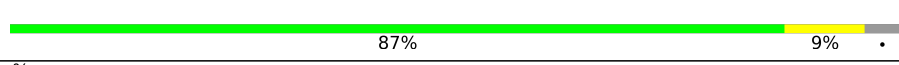

The reported resolution of this entry is 2.00 Å.

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)
RSRZ outliers	127900	7900 (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  $\leq 5\%$ . 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	401	 87% 9% .
1	B	401	 88% 7% .

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5806 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 Acetyl-CoA C-acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	386	Total 2858	C 1804	N 489	O 552	S 13	0	1	0
1	B	384	Total 2842	C 1798	N 484	O 547	S 13	0	1	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	394	LEU	-	expression tag	UNP B9DX69
A	395	GLU	-	expression tag	UNP B9DX69
A	396	HIS	-	expression tag	UNP B9DX69
A	397	HIS	-	expression tag	UNP B9DX69
A	398	HIS	-	expression tag	UNP B9DX69
A	399	HIS	-	expression tag	UNP B9DX69
A	400	HIS	-	expression tag	UNP B9DX69
A	401	HIS	-	expression tag	UNP B9DX69
B	394	LEU	-	expression tag	UNP B9DX69
B	395	GLU	-	expression tag	UNP B9DX69
B	396	HIS	-	expression tag	UNP B9DX69
B	397	HIS	-	expression tag	UNP B9DX69
B	398	HIS	-	expression tag	UNP B9DX69
B	399	HIS	-	expression tag	UNP B9DX69
B	400	HIS	-	expression tag	UNP B9DX69
B	401	HIS	-	expression tag	UNP B9DX69


- Molecule 2 is water.

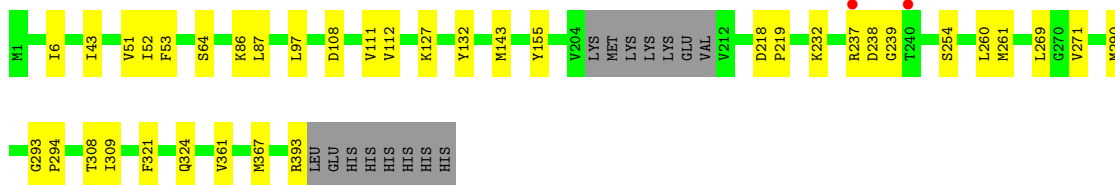
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	52	Total 52	O 52	0	0
2	B	54	Total 54	O 54	0	0

### 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 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.

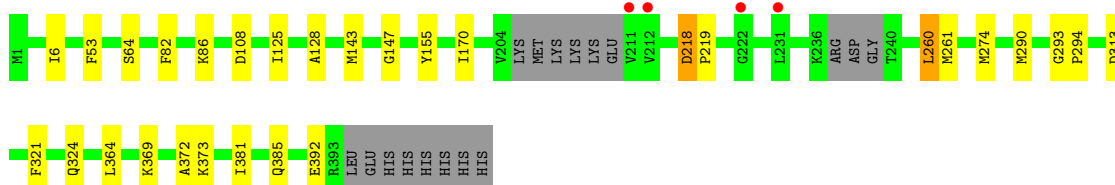
- Molecule 1: Acetyl-CoA C-acetyltransferase

Chain A:  87% 9%



- Molecule 1: Acetyl-CoA C-acetyltransferase

Chain B:  88% 7%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.57Å 73.57Å 274.00Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.65 – 2.00 46.65 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.8 (46.65-2.00) 97.8 (46.65-2.00)	Depositor EDS
$R_{merge}$	0.03	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.70 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.8.0411	Depositor
R, $R_{free}$	0.209 , 0.240 0.218 , 0.244	Depositor DCC
$R_{free}$ test set	2011 reflections (3.46%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.3	Xtrriage
Anisotropy	0.170	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 37.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	0.054 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5806	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.08% 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	0.41	0/2897	0.72	0/3913
1	B	0.41	0/2880	0.69	0/3890
All	All	0.41	0/5777	0.71	0/7803

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	2858	0	2915	18	0
1	B	2842	0	2905	16	0
2	A	52	0	0	1	0
2	B	54	0	0	4	0
All	All	5806	0	5820	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.

The worst 5 of 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:A:43:ILE:HD11	1:A:269:LEU:HD13	1.57	0.84

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:6:ILE:HG12	1:A:260:LEU:HD22	1.78	0.65
1:A:86:LYS:NZ	2:A:501:HOH:O	2.32	0.61
1:A:308:THR:HG22	1:A:309:ILE:H	1.65	0.61
1:B:369:LYS:HE2	2:B:518:HOH:O	2.00	0.61

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	383/401 (96%)	368 (96%)	13 (3%)	2 (0%)	29	23
1	B	379/401 (94%)	362 (96%)	15 (4%)	2 (0%)	29	23
All	All	762/802 (95%)	730 (96%)	28 (4%)	4 (0%)	29	23

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	64	SER
1	B	64	SER
1	A	238	ASP
1	B	147	GLY

### 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	297/311 (96%)	289 (97%)	8 (3%)	44	46
1	B	296/311 (95%)	289 (98%)	7 (2%)	49	51
All	All	593/622 (95%)	578 (98%)	15 (2%)	47	49

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	361	VAL
1	B	260	LEU
1	B	53	PHE
1	B	290	MET
1	B	170	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	68	GLN
1	B	104	ASN
1	B	324	GLN
1	B	252	ASN
1	A	324	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 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 [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	386/401 (96%)	-0.22	2 (0%) 91 90	22, 30, 48, 86	0
1	B	384/401 (95%)	-0.23	4 (1%) 82 81	21, 31, 49, 62	0
All	All	770/802 (96%)	-0.23	6 (0%) 86 85	21, 31, 48, 86	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	240	THR	4.7
1	A	237	ARG	2.7
1	B	212	VAL	2.2
1	B	211	VAL	2.1
1	B	231	LEU	2.1

### 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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.