



# wwPDB X-ray Structure Validation Summary Report

Sep 24, 2023 – 06:22 PM EDT

PDB ID : 5UPB  
Title : Swit\_4259, an Acetoacetate Decarboxylase-like Enzyme from *Sphingomonas wittichii* RW1  
Authors : Mydy, L.S.; Silvaggi, N.R.  
Deposited on : 2017-02-02  
Resolution : 1.91 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 : 2.35.1  
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) : 2.35.1

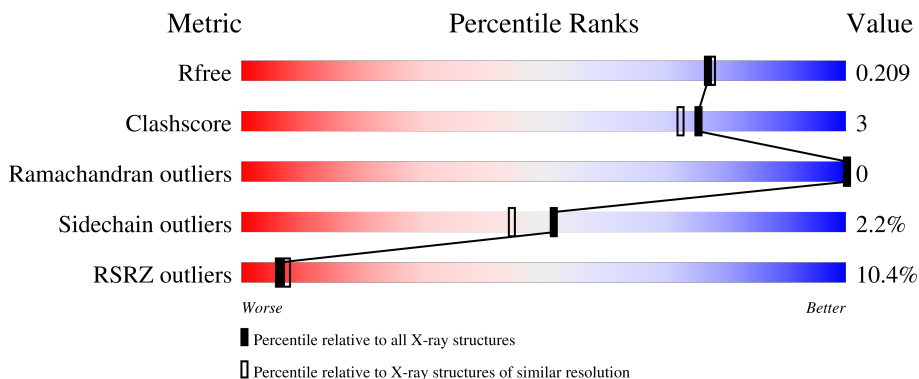
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.91 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	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	258	 5% 90% 9%
1	B	258	 6% 93% 6%
1	C	258	 9% 88% 11%
1	D	258	 20% 89% 9%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 15566 atoms, of which 7386 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 Acetoacetate decarboxylase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	256	3798	1239	1852	344	355	8	0	0	0
1	B	256	3813	1239	1867	344	355	8	0	0	0
1	C	255	3742	1233	1807	342	352	8	0	0	0
1	D	254	3785	1227	1860	341	349	8	0	0	0

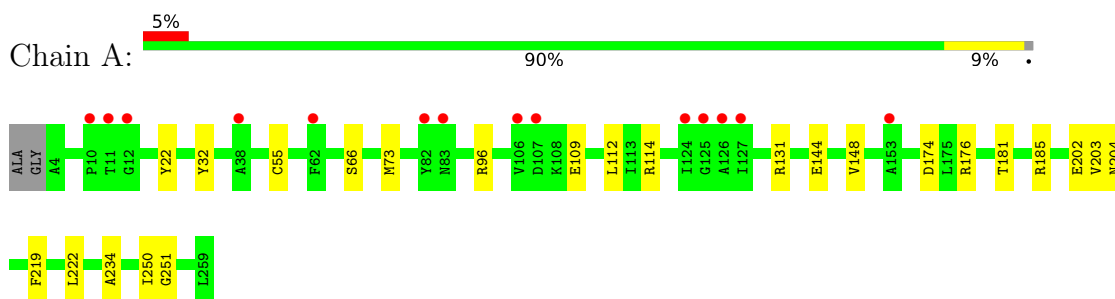
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	118	Total 118	O 118	0	0
2	B	116	Total 116	O 116	0	0
2	C	108	Total 108	O 108	0	0
2	D	86	Total 86	O 86	0	0

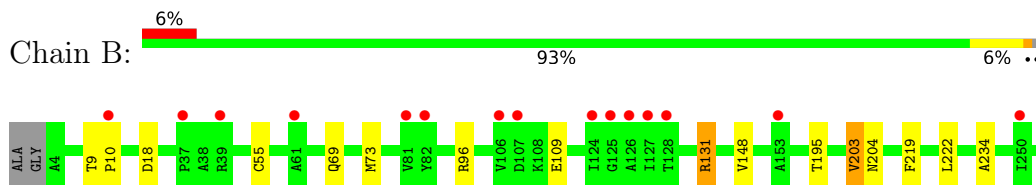
### 3 Residue-property plots

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.

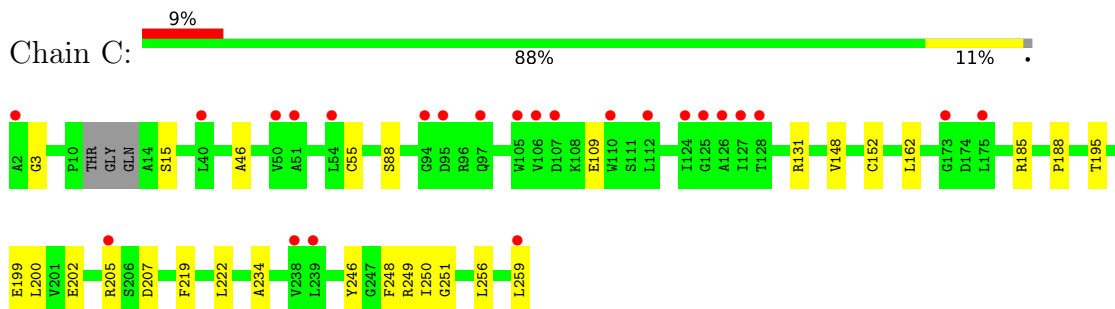
- Molecule 1: Acetoacetate decarboxylase



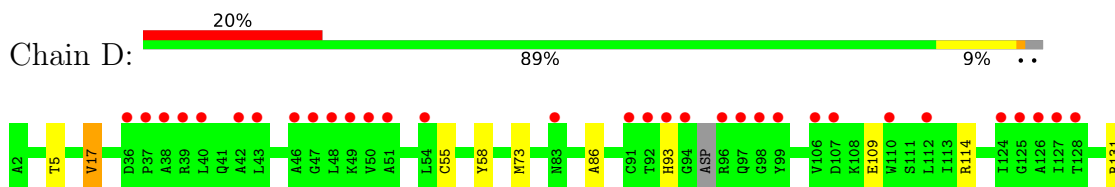
- Molecule 1: Acetoacetate decarboxylase

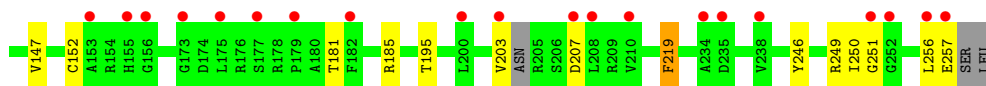


- Molecule 1: Acetoacetate decarboxylase



- Molecule 1: Acetoacetate decarboxylase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.04Å 75.04Å 179.85Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.06 – 1.91 44.06 – 1.91	Depositor EDS
% Data completeness (in resolution range)	99.1 (44.06-1.91) 99.9 (44.06-1.91)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.29 (at 1.91Å)	Xtriage
Refinement program	PHENIX 1.11_2567	Depositor
R, $R_{free}$	0.172 , 0.201 0.177 , 0.209	Depositor DCC
$R_{free}$ test set	3021 reflections (3.45%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.5	Xtriage
Anisotropy	0.540	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 47.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.013 for -h,-k,l 0.052 for h,-h-k,-l 0.028 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	15566	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.12% 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	0.58	1/2000 (0.1%)	0.79	2/2729 (0.1%)
1	B	0.60	1/2000 (0.1%)	0.78	3/2729 (0.1%)
1	C	0.59	2/1988 (0.1%)	0.72	0/2711
1	D	0.57	2/1977 (0.1%)	0.77	4/2694 (0.1%)
All	All	0.58	6/7965 (0.1%)	0.77	9/10863 (0.1%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	55	CYS	CB-SG	-8.92	1.67	1.82
1	C	55	CYS	CB-SG	-8.07	1.68	1.82
1	D	55	CYS	CB-SG	-7.49	1.69	1.82
1	A	55	CYS	CB-SG	-7.12	1.70	1.82
1	D	152	CYS	CB-SG	-6.82	1.70	1.82

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	114	ARG	NE-CZ-NH1	9.06	124.83	120.30
1	A	114	ARG	NE-CZ-NH2	-8.82	115.89	120.30
1	B	55	CYS	CB-CA-C	-6.63	97.14	110.40
1	D	219	PHE	CB-CG-CD2	-6.38	116.33	120.80
1	D	219	PHE	CB-CG-CD1	6.03	125.02	120.80

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	1946	1852	1892	15	0
1	B	1946	1867	1892	8	0
1	C	1935	1807	1881	16	0
1	D	1925	1860	1872	13	0
2	A	118	0	0	3	0
2	B	116	0	0	0	0
2	C	108	0	0	0	1
2	D	86	0	0	0	1
All	All	8180	7386	7537	47	1

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 47 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:96:ARG:NE	1:A:202:GLU:OE2	2.07	0.87
1:C:185:ARG:NH1	1:D:195:THR:O	2.13	0.77
1:A:22:TYR:CE1	2:A:302:HOH:O	2.37	0.76
1:A:22:TYR:CZ	2:A:302:HOH:O	2.42	0.71
1:A:181:THR:HB	1:A:203:VAL:CG1	2.28	0.64

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:366:HOH:O	2:D:370:HOH:O[3_555]	2.08	0.12

## 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	254/258 (98%)	252 (99%)	2 (1%)	0	100	100
1	B	254/258 (98%)	251 (99%)	3 (1%)	0	100	100
1	C	251/258 (97%)	250 (100%)	1 (0%)	0	100	100
1	D	248/258 (96%)	245 (99%)	3 (1%)	0	100	100
All	All	1007/1032 (98%)	998 (99%)	9 (1%)	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	195/195 (100%)	192 (98%)	3 (2%)	65	61
1	B	195/195 (100%)	191 (98%)	4 (2%)	53	46
1	C	193/195 (99%)	188 (97%)	5 (3%)	46	37
1	D	191/195 (98%)	186 (97%)	5 (3%)	46	37
All	All	774/780 (99%)	757 (98%)	17 (2%)	52	45

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	109	GLU
1	D	246	TYR
1	C	88	SER
1	C	109	GLU
1	C	162	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	186	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](#)

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	256/258 (99%)	0.24	14 (5%) 25 28	25, 38, 81, 125	0
1	B	256/258 (99%)	0.31	16 (6%) 20 22	25, 38, 88, 116	0
1	C	255/258 (98%)	0.56	24 (9%) 8 10	27, 43, 83, 202	0
1	D	254/258 (98%)	0.82	52 (20%) 1 1	30, 49, 94, 122	0
All	All	1021/1032 (98%)	0.48	106 (10%) 6 7	25, 41, 86, 202	0

The worst 5 of 106 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	95	ASP	4.4
1	C	124	ILE	4.4
1	D	92	THR	4.0
1	D	175	LEU	3.8
1	A	124	ILE	3.8

### 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.