



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 17, 2016 – 10:35 AM EDT

PDB ID : 5I6G  
Title : Crystal structure of C-terminal variant 2 of Chaetomium thermophilum acetyl-CoA carboxylase  
Authors : Hunkeler, M.; Stuttfeld, E.; Hagmann, A.; Imseng, S.; Maier, T.  
Deposited on : 2016-02-16  
Resolution : 4.50 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  
<http://wwpdb.org/validation/2016/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 : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027257  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027257

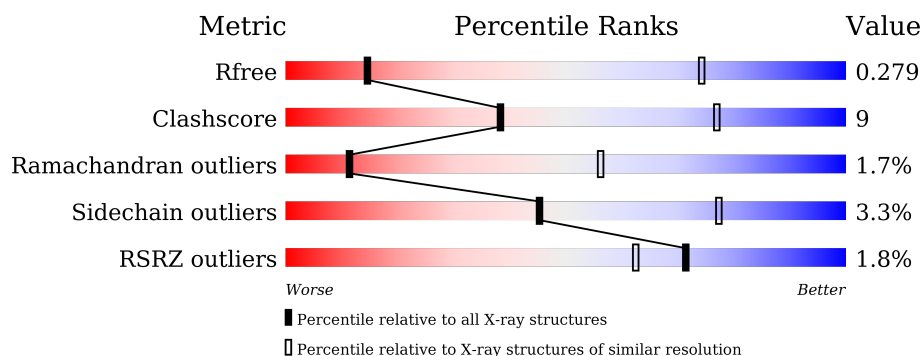
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 4.50 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))
$R_{free}$	91344	1071 (5.40-3.60)
Clashscore	102246	1003 (5.30-3.62)
Ramachandran outliers	100387	1117 (5.40-3.60)
Sidechain outliers	100360	1099 (5.40-3.60)
RSRZ outliers	91569	1075 (5.40-3.60)

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. 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	1161	<div> <div> <div></div> <div>66%</div> <div>21%</div> <div>•</div> <div>11%</div> </div> </div>
1	B	1161	<div> <div> <div>2%</div> <div>68%</div> <div>20%</div> <div>•</div> <div>11%</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 16405 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 carboxylase-like protein, Acetyl-CoA carboxylase-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1028	Total	C	N	O	S	0	0	0
			8193	5204	1434	1526	29			
1	B	1030	Total	C	N	O	S	0	0	0
			8212	5216	1439	1528	29			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1113	GLY	-	expression tag	UNP G0S3L5
B	1113	GLY	-	expression tag	UNP G0S3L5





GLY	ASN	LEU	ARG	VAL	ARG	TYR	HIS	ASP	GLU	ARG	PRO	TYR	PHE	ASP	TRP	ASP	PHE	ALA	LEU	ARG	LYS	SER	GLY	ALA	ASN	GLN	THR	GLU	SER	LEU	SER	MET	HIS	LYS	GLU	THR	LYS	VAL	PRO	SER	SER	PRO	ALA	THR	PRO	VAL	GLU	ASN	ASP	PHE	LYS	ARG	ILE	HIS	SER	ILE	SER																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.17Å 153.45Å 249.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.10 – 4.50 49.10 – 4.50	Depositor EDS
% Data completeness (in resolution range)	99.4 (49.10-4.50) 99.4 (49.10-4.50)	Depositor EDS
$R_{merge}$	0.27	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.38 (at 4.45Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.236 , 0.240 0.274 , 0.279	Depositor DCC
$R_{free}$ test set	1112 reflections (4.77%)	DCC
Wilson B-factor (Å <sup>2</sup> )	202.9	Xtriage
Anisotropy	0.442	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 255.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	16405	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	275.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.64% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.51	0/8311	0.70	1/11255 (0.0%)
1	B	0.49	0/8330	0.68	0/11280
All	All	0.50	0/16641	0.69	1/22535 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1506	TYR	CA-CB-CG	5.16	123.20	113.40

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	8193	0	8096	152	0
1	B	8212	0	8120	157	0
All	All	16405	0	16216	296	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 9.

The worst 5 of 296 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:1515:SER:HA	1:B:1599:SER:O	1.25	1.28
1:B:1319:ASP:HB2	1:B:1344:PRO:HG2	1.39	1.02
1:B:1515:SER:CA	1:B:1599:SER:O	2.19	0.89
1:A:1336:ARG:HB3	1:A:1525:TYR:CD2	2.09	0.88
1:B:1609:MET:HE1	1:B:1630:VAL:CG1	2.06	0.86

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	1006/1161 (87%)	898 (89%)	89 (9%)	19 (2%)	10	54
1	B	1008/1161 (87%)	897 (89%)	96 (10%)	15 (2%)	13	58
All	All	2014/2322 (87%)	1795 (89%)	185 (9%)	34 (2%)	11	56

5 of 34 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1864	SER
1	A	2225	ASN
1	B	1864	SER
1	A	1333	GLU
1	A	1483	SER

### 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	872/990 (88%)	838 (96%)	34 (4%)	39	73
1	B	874/990 (88%)	851 (97%)	23 (3%)	54	81
All	All	1746/1980 (88%)	1689 (97%)	57 (3%)	45	77

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2088	LEU
1	A	2196	ASN
1	B	2151	ARG
1	A	2106	LYS
1	A	2151	ARG

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

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	2261:ARG	C	2300:UNK	N	11.76

## 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	1016/1161 (87%)	-0.11	10 (0%) 84 78	129, 237, 392, 485	0
1	B	1018/1161 (87%)	-0.00	27 (2%) 58 48	147, 289, 446, 500	0
All	All	2034/2322 (87%)	-0.06	37 (1%) 71 62	129, 257, 429, 500	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1476	ARG	5.7
1	B	1297	CYS	4.9
1	B	1298	GLY	4.6
1	B	1371	ARG	3.9
1	B	1296	ILE	3.6

### 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 carbohydrates 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.