



# wwPDB X-ray Structure Validation Summary Report

Jun 22, 2024 – 11:52 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

<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 : 2.37.1  
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) : 2.37.1

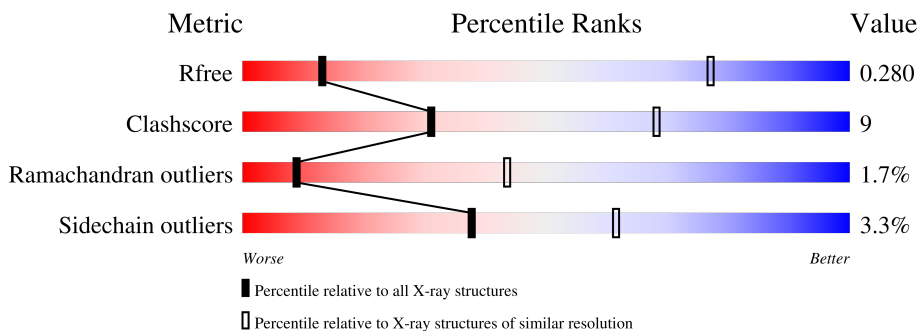
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 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))
$R_{free}$	130704	1055 (5.20-3.80)
Clashscore	141614	1123 (5.20-3.80)
Ramachandran outliers	138981	1069 (5.20-3.80)
Sidechain outliers	138945	1050 (5.20-3.80)

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	1161	
1	B	1161	

## 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
			Total	C	N	O	S			
1	A	1028	8193	5204	1434	1526	29	0	0	0
1	B	1030	8212	5216	1439	1528	29	0	0	0

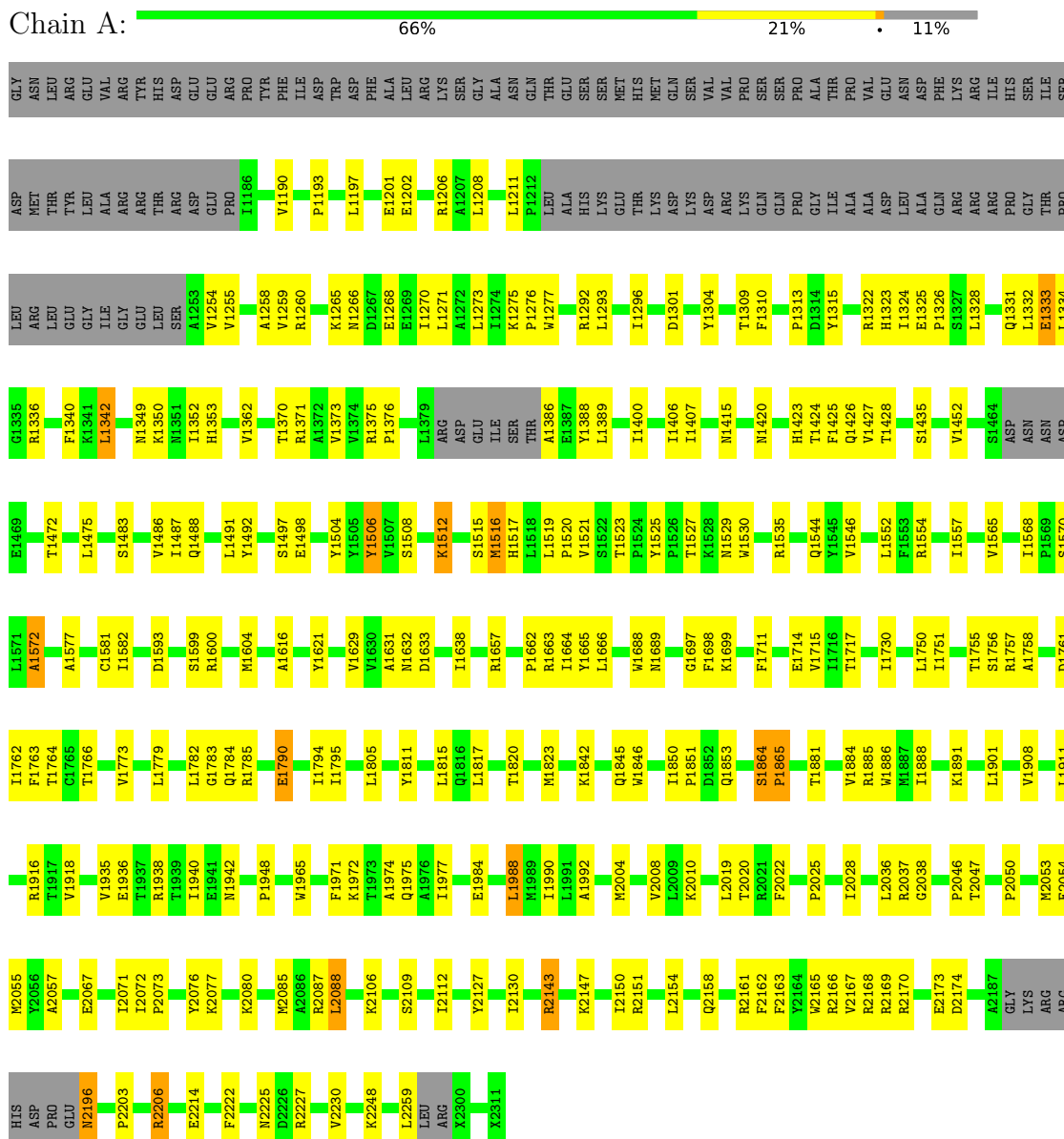
There are 2 discrepancies between the modelled and reference sequences:

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

### 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: Acetyl-CoA carboxylase-like protein, Acetyl-CoA carboxylase-like protein



- Molecule 1: Acetyl-CoA carboxylase-like protein, Acetyl-CoA carboxylase-like protein



R2226	M2085	I1940	G1783	A1816	S1483	K1341	ARG
R2227	A2086	A1961	Q1784	Y1621	V1486	P1344	LEU
V2230	R2087	L2088	R1785	V1629	E1493	N1349	GLU
E2242	S2109	W1965	E1790	V1630	E1494	K1350	ILE
K2246	F1971	F1971	I1794	A1631	E1497	W1353	GLY
I2252	T1973	T1973	I1795	N1632	E1498	I1352	LEU
Q2255	A1974	A1974	L1805	D1633	E1498	Y1355	SER
X2311	A1976	A1976	I1815	I1638	Y1504	R1360	A1283
	I1977	I1977	Q1816	F1649	I1513	V1254	V1259
	L1988	L1988	L1817	R1657	I1513	R1280	R1280
	M1989	M1989	T1820	P1662	G1514	D1261	D1261
	D2139	I1990	T1820	P1662	S1515	A1262	A1262
	R2140	L1991	M1823	R1663	M1516	E1263	E1263
	A1992	A1992	M1823	R1664	I1519	G1264	G1264
	R2143	Q2001	K1842	I1664	P1520	K1265	K1265
	K2147	Q2001	Q1845	I1665	V1521	M1286	M1286
	I2150	M2004	Q1846	L1666	Y1525	D1287	D1287
	R2151	V2008	W1846	S1670	Y1525	E1288	E1288
	L2154	L2019	P1851	L1674	W1530	I1270	I1270
	Q2158	T2020	D1882	V1686	L1531	L1271	L1271
	R2161	R2021	R1854	A1687	Q1532	A1272	A1272
	F2162	F2022	S1864	W1688	F1533	L1273	L1273
	F2163	R2022	P1865	A1687	A1403	W1277	W1277
	W2165	P2025	P1876	K1699	I1406	R1292	R1292
	R2166	T2028	P1877	F1702	I1407	F1295	F1295
	V2167	L2036	T1881	Y1702	Y1445	I1296	I1296
	R2168	R2037	V1884	F1711	F1552	R1299	R1299
	R2169	G2038	R1885	E1714	R1554	HIS	HIS
	R2170	P2046	W1886	E1714	I1557	MET	MET
	E2173	T2047	M1887	K1729	I1568	ASP	ASP
	D2174	P2050	I1888	L1750	S1570	LYS	LYS
	A2187	P2050	K1891	I1751	L1571	LYS	LYS
	GLY	M2053	L1901	T1755	A1572	ASP	ASP
	LYS	E2054	L1901	T1755	A1577	ASP	ASP
	ARG	M2055	V1908	S1756	C1581	ARG	ARG
	HIS	A2057	E1909	R1757	I1582	VAL	VAL
	ASP	E2067	T1910	A1758	M1585	GLN	GLN
	PRO	T2072	L1911	F1763	M1585	PRO	PRO
	GLU	P2073	R1916	T1766	D1593	ALA	ALA
	N2196	P2073	T1917	T1766	D1593	ALA	ALA
	P2203	Y2076	V1918	V1773	S1599	ASP	ASP
	R2206	K2077	V1985	L1779	P1602	ASP	ASP
	N2225	K2080	R1938	V1780	M1609	ASP	ASP
			T1939	L1781	L1475	THR	THR
				L1782	R1476	PRO	PRO
						LEU	LEU
						SER	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)	Xtrriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.236 , 0.240 0.279 , 0.280	Depositor DCC
$R_{free}$ test set	1112 reflections (4.76%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	202.9	Xtrriage
Anisotropy	0.442	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 256.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.22$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	16405	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	275.0	wwPDB-VP

Xtrriage'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, 0.333 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	0.51	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%)	8	41
1	B	1008/1161 (87%)	897 (89%)	96 (10%)	15 (2%)	10	46
All	All	2014/2322 (87%)	1795 (89%)	185 (9%)	34 (2%)	9	43

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%)	32	57
1	B	874/990 (88%)	851 (97%)	23 (3%)	46	67
All	All	1746/1980 (88%)	1689 (97%)	57 (3%)	38	61

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

Mol	Chain	Res	Type
1	A	2151	ARG
1	B	2196	ASN
1	B	1301	ASP
1	B	2174	ASP
1	B	2088	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

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

### 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.