



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 17, 2016 – 09:32 AM EDT

PDB ID : 5I6H
Title : Crystal structure of CD-CT domains of Chaetomium thermophilum acetyl-CoA carboxylase
Authors : Hunkeler, M.; Stuttfeld, E.; Hagmann, A.; Imseng, S.; Maier, T.
Deposited on : 2016-02-16
Resolution : 7.20 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
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

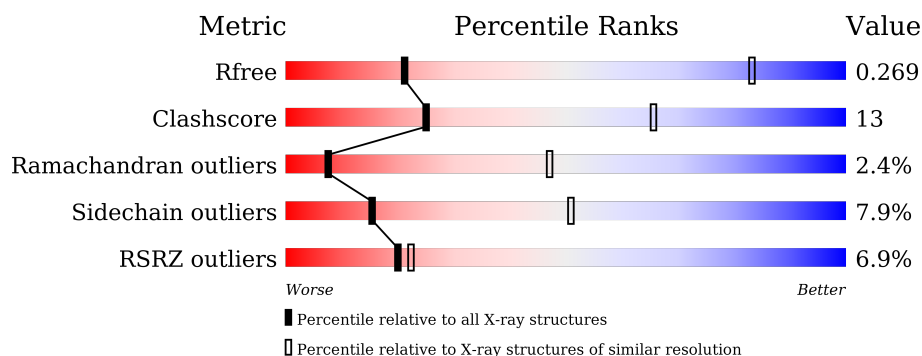
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 7.20 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	1014 (9.50-3.66)
Clashscore	102246	1063 (10.00-3.70)
Ramachandran outliers	100387	1035 (9.50-3.66)
Sidechain outliers	100360	1005 (9.50-3.66)
RSRZ outliers	91569	1013 (9.50-3.66)

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 ≥ 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	1487	<div> <div>7%</div> <div>63%</div> <div>28%</div> <div>• 6%</div> </div>
1	B	1487	<div> <div>6%</div> <div>64%</div> <div>27%</div> <div>• 5%</div> </div>

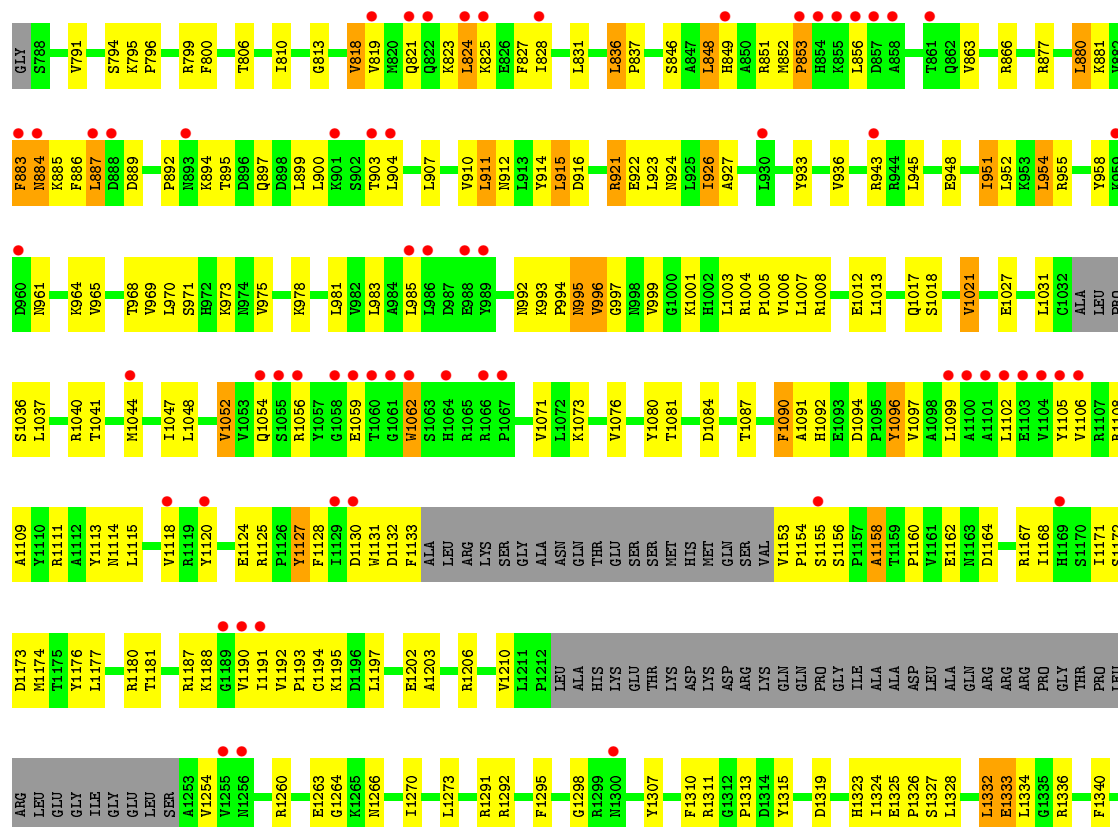
2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 22543 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1404	Total	C	N	O	S	0	0	0
			11262	7144	1983	2096	39			
1	B	1406	Total	C	N	O	S	0	0	0
			11281	7156	1988	2098	39			



L2199	L2071	Q1975	V1846	T1717	Q1575	ASN	P1344
A2200	I2072	D1979	L1847	H1728	P1576	ASN	M1349
S2202	P2073	F1980	P1851	K1729	V1588	ASP	
P2203	Y2076	N1981	D1852	E1737	D1593	E1469	I1352
H2209	K2080	Q1985	Q1853	E1738	V1477		H1353
V2211	L2088	L1988	R1854	G1739	I1478		V1355
Q2212	L2088	M1989	L1858	L1740	V1598		
L2213	S2109	I1990	P1863	G1741	S1599		F1369
	I2112	L1991	P1864	V1742	R1600		T1370
W2216	A1992	A1992	S1865	E1743	M1604		R1371
G2217	N1993	N1993	P1866	C1744	S1483		
G2218	F1997	F1997	D1866	L1745	G1484		
L2219			P1867	L1750	M1609		R1375
N2225	I2130	Q2001	V1884	L1750	F1485		
	F2134	M2004	R1885	S1756	V1486		R1378
K2243	L2137	E2007	R1886	R1757	Y1630		L1379
L2244	R2138	V2008	M1887	A1758	A1631		ARG
E2245	D2139		I1888	Y1759	N1632		ASP
K2246	R2140		K1891	N1760	D1633		GLU
	R2143	A2018	Q1898	D1761	I1633		ILE
K2249	K2147	L2019	L1901	I1762	S1640		SER
	F2022	R2021	V1908	F1763	L1496		THR
L2252	I2150	E2023	E1909		W1502		A1386
M2256	R2151	K2024	A1915	R1771	S1639		E1387
	R2166	P2025	R1916	S1772	Y1503		Y1388
	V2167	I2028	T1917	V1773	Y1504		L1389
	R2168	Y2029	V1918	G1776	Y1505		I1390
	R2169	P2032	V1919		Y1506		
	R2170	H2033	G1921	V1780	C1652		D1402
L2171	L2171	L2036	R1922	R1781	G1514		A1403
E2172	E2172	R2037	A1923	S1772	S1515		I1406
D2173	E2173	G2038	I1928	V1773	M1516		N1415
D2174	D2174		P1929	G1776	H1517		H1416
D2175	L2175		M1930		L1518		
L2177	L2177	P2046	V1935	V1787	L1519		
	L2180	T2047		E1790	P1520		S1422
		P2050	I1940	L1795	V1521		H1423
			E1941		S1522		T1424
A2187	GLY	M2053	N1942	S1813	F1675		F1425
LYS	LYS	E2054	M2055	L1815	L1676		Q1426
ARG	ARG	Y2056	T1958	Q1816	W1688		
ARG	ARG	A2057		L1817	P1694		W1448
HIS	HIS		W1965		F1698		R1451
ASP	ASP	V2065			Y1702		V1452
PRO	PRO	L2066	N1968	T1820	E1563		H1453
GLU	GLU	E2067		K1842	Y1545		Q1454
		P2068	F1971	I1843	V1546		
N2196	N2196		K1972	E1714	I1557		
T2197	T2197	E2069		V1715			
S2199	S2199	C2070		I1716			
							ASP

4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	295.02Å 295.02Å 189.52Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.85 – 7.20 49.85 – 7.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.85-7.20) 100.0 (49.85-7.20)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.72 (at 7.37Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.232 , 0.248 0.260 , 0.269	Depositor DCC
R_{free} test set	696 reflections (4.97%)	DCC
Wilson B-factor (Å ²)	662.8	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.23 , 411.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.021 for -h,-k,l	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	22543	wwPDB-VP
Average B, all atoms (Å ²)	271.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.82% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.49	0/11439	0.72	3/15486 (0.0%)
1	B	0.50	0/11458	0.71	0/15511
All	All	0.49	0/22897	0.72	3/30997 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1988	LEU	CB-CG-CD2	5.41	120.20	111.00
1	A	1988	LEU	CB-CG-CD1	5.17	119.78	111.00
1	A	1090	PHE	C-N-CA	5.09	134.41	121.70

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	11262	0	11181	289	0
1	B	11281	0	11205	295	0
All	All	22543	0	22386	571	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 13.

The worst 5 of 571 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:1389:LEU:HD11	1:B:1425:PHE:HB3	1.33	1.08
1:A:994:PRO:HA	1:A:1037:LEU:HB3	1.35	1.07
1:B:994:PRO:HA	1:B:1037:LEU:HB3	1.40	1.01
1:A:1090:PHE:HB2	1:A:1193:PRO:HB3	1.39	0.99
1:B:1176:TYR:HE2	1:B:1291:ARG:HD3	1.23	0.98

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	1378/1487 (93%)	1221 (89%)	121 (9%)	36 (3%)	7	45
1	B	1380/1487 (93%)	1221 (88%)	130 (9%)	29 (2%)	9	50
All	All	2758/2974 (93%)	2442 (88%)	251 (9%)	65 (2%)	7	47

5 of 65 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	819	VAL
1	A	943	ARG
1	A	1092	HIS
1	A	1117	GLU
1	A	1128	PHE

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	1213/1282 (95%)	1116 (92%)	97 (8%)	15	50
1	B	1215/1282 (95%)	1119 (92%)	96 (8%)	15	51
All	All	2428/2564 (95%)	2235 (92%)	193 (8%)	15	51

5 of 193 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2140	ARG
1	B	916	ASP
1	B	2022	PHE
1	A	2151	ARG
1	B	818	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	995	ASN
1	B	1043	GLN
1	B	1460	ASN

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 [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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	9.73

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, 95th 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 ²)	Q<0.9
1	A	1392/1487 (93%)	0.15	102 (7%) 18 21	33, 251, 497, 500	0
1	B	1394/1487 (93%)	0.09	89 (6%) 23 24	13, 262, 498, 500	0
All	All	2786/2974 (93%)	0.12	191 (6%) 20 22	13, 258, 497, 500	0

The worst 5 of 191 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1353	HIS	10.5
1	A	1347	THR	8.4
1	A	1301	ASP	8.3
1	B	1103	GLU	7.9
1	A	959	LYS	7.8

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 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.