WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Apr 21, 2024 – 05:08 pm BST

PDB ID	:	2C7I
Title	:	Structure of protein Ta0514, putative lipoate protein ligase from T. acidophilum.
Authors	:	Mcmanus, E.; Perham, R.N.; Luisi, B.F.
Deposited on	:	2005-11-24
Resolution	:	2.10 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

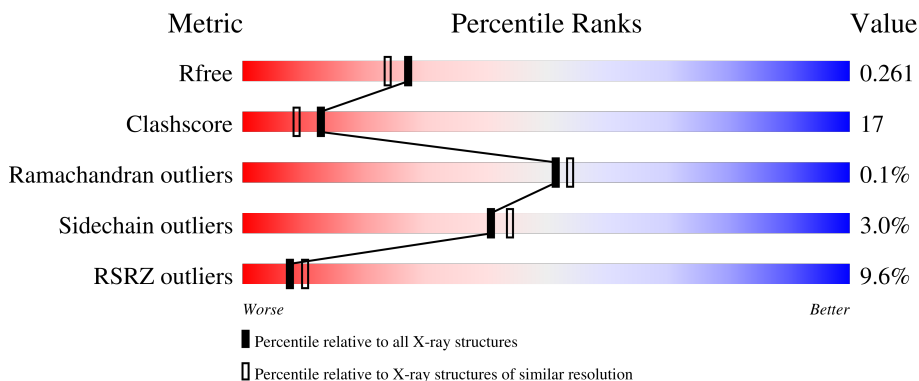|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.10 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	Percentile Ranks	Value
Rfree		0.261
Clashscore		17
Ramachandran outliers		0.1%
Sidechain outliers		3.0%
RSRZ outliers		9.6%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 >=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 <=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	262	5% 70% 19% • 7%
1	B	262	5% 74% 16% • 8%
1	C	262	10% 61% 27% • • 8%
1	D	262	16% 59% 29% • 8%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 7610 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 PUTATIVE LIPOATE PROTEIN LIGASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	244	Total	C	N	O	S	9	2	0
			1921	1203	335	370	13			
1	B	240	Total	C	N	O	S	17	1	0
			1861	1169	318	361	13			
1	C	240	Total	C	N	O	S	65	0	0
			1820	1148	309	351	12			
1	D	240	Total	C	N	O	S	61	0	0
			1820	1147	315	346	12			

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	82	Total	O	0	0
			82	82		
2	B	59	Total	O	0	0
			59	59		
2	C	27	Total	O	0	0
			27	27		
2	D	20	Total	O	0	0
			20	20		

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.
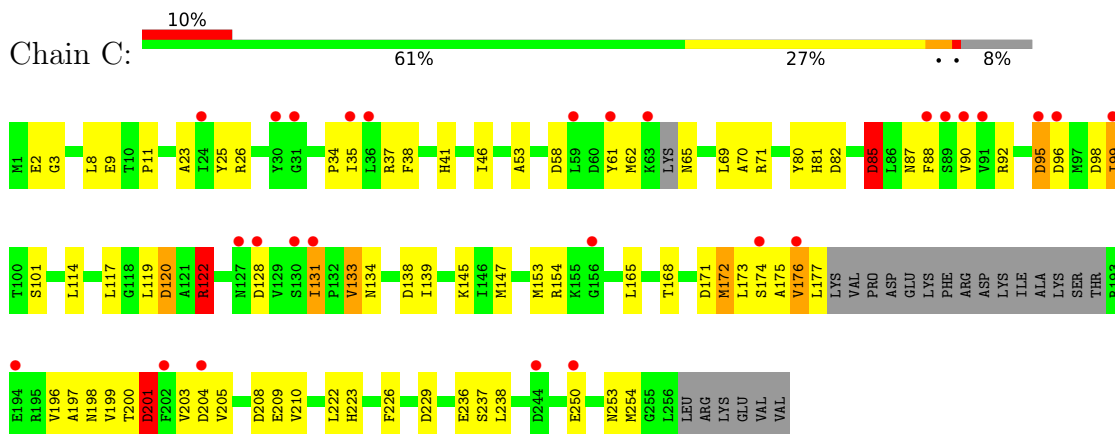
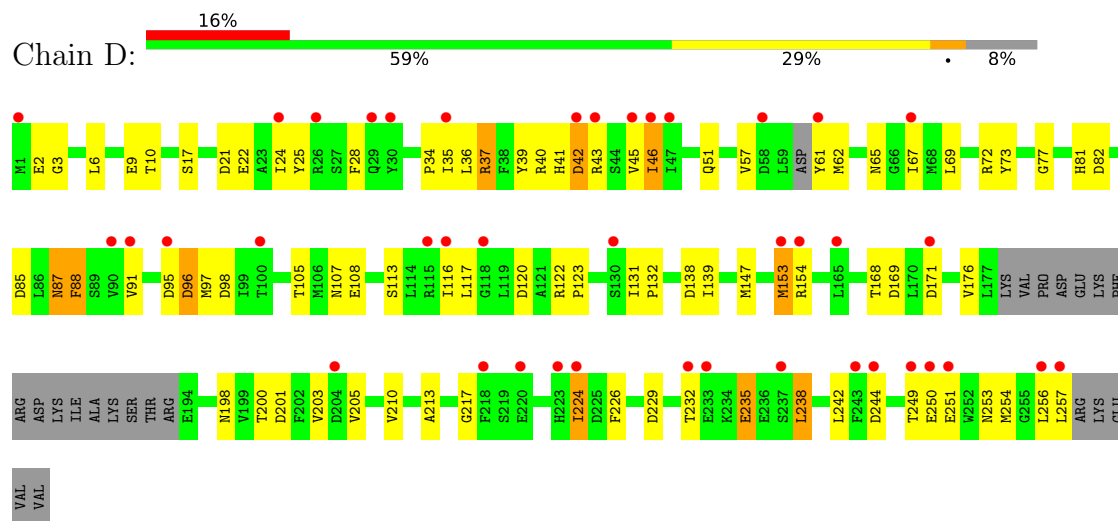• Molecule 1: PUTATIVE LIPOATE PROTEIN LIGASE
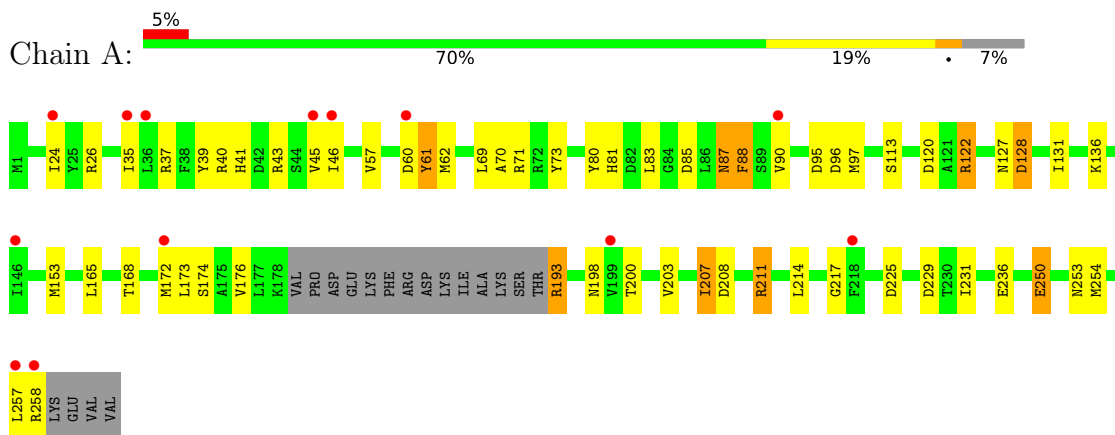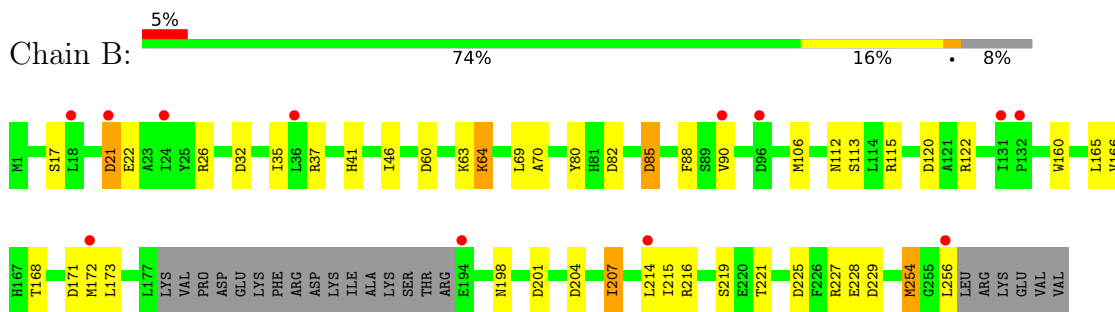
Chain A:

• Molecule 1: PUTATIVE LIPOATE PROTEIN LIGASE

Chain B:

• Molecule 1: PUTATIVE LIPOATE PROTEIN LIGASE

Chain C:

• Molecule 1: PUTATIVE LIPOATE PROTEIN LIGASE

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	53.23Å 118.36Å 105.59Å 90.00° 93.75° 90.00°	Depositor
Resolution (Å)	19.87 – 2.10 19.85 – 2.10	Depositor EDS
% Data completeness (in resolution range)	94.8 (19.87-2.10) 94.8 (19.85-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.52 (at 2.09Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.208 , 0.252 0.221 , 0.261	Depositor DCC
R_{free} test set	3634 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	38.9	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 67.4	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7610	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	0.60	1/1956 (0.1%)	1.06	14/2637 (0.5%)
1	B	0.58	1/1893 (0.1%)	0.80	10/2553 (0.4%)
1	C	0.45	2/1848 (0.1%)	1.48	20/2500 (0.8%)
1	D	0.46	2/1847 (0.1%)	0.84	15/2495 (0.6%)
All	All	0.53	6/7544 (0.1%)	1.08	59/10185 (0.6%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	C	0	5
1	D	0	1
All	All	0	8

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	211	ARG	CZ-NH1	10.59	1.46	1.33
1	B	122	ARG	NE-CZ	-9.63	1.20	1.33
1	D	37	ARG	C-N	-7.83	1.16	1.34
1	D	235	GLU	CB-CG	7.49	1.66	1.52
1	C	122	ARG	CZ-NH2	6.74	1.41	1.33

The worst 5 of 59 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	122	ARG	NE-CZ-NH2	-57.69	91.45	120.30
1	A	211	ARG	NE-CZ-NH1	-29.12	105.74	120.30
1	C	85	ASP	CB-CG-OD1	-16.06	103.85	118.30
1	C	122	ARG	NH1-CZ-NH2	14.94	135.84	119.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	36	LEU	O-C-N	-13.23	101.53	122.70

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	122	ARG	Sidechain
1	A	211	ARG	Sidechain
1	C	101	SER	Mainchain
1	C	85	ASP	Sidechain
1	C	95	ASP	Mainchain

5.2 Too-close contacts ⓘ

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	1921	0	1863	59	0
1	B	1861	0	1794	30	0
1	C	1820	0	1726	57	0
1	D	1820	0	1739	89	0
2	A	82	0	0	2	0
2	B	59	0	0	2	0
2	C	27	0	0	0	0
2	D	20	0	0	1	0
All	All	7610	0	7122	233	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 17.

The worst 5 of 233 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:D:45:VAL:HG11	1:D:57:VAL:CG2	1.84	1.07
1:A:45:VAL:CG2	1:A:176:VAL:HG11	1.87	1.05
1:C:131:ILE:H	1:C:131:ILE:HD13	0.89	1.05

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:131:ILE:H	1:C:131:ILE:CD1	1.65	1.04
1:D:238:LEU:O	1:D:238:LEU:HD22	1.58	1.03

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	242/262 (92%)	236 (98%)	6 (2%)	0	100	100
1	B	237/262 (90%)	226 (95%)	11 (5%)	0	100	100
1	C	234/262 (89%)	220 (94%)	13 (6%)	1 (0%)	34	32
1	D	234/262 (89%)	219 (94%)	15 (6%)	0	100	100
All	All	947/1048 (90%)	901 (95%)	45 (5%)	1 (0%)	51	54

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	201	ASP

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	201/226 (89%)	195 (97%)	6 (3%)	41	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	193/226 (85%)	188 (97%)	5 (3%)	46	50
1	C	183/226 (81%)	176 (96%)	7 (4%)	33	34
1	D	183/226 (81%)	178 (97%)	5 (3%)	44	48
All	All	760/904 (84%)	737 (97%)	23 (3%)	41	44

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	133	VAL
1	C	237	SER
1	C	176	VAL
1	D	46	ILE
1	B	64	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	65	ASN
1	C	198	ASN
1	D	198	ASN
1	C	253	ASN
1	B	41	HIS

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 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	37:ARG	C	38:PHE	N	1.16

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	244/262 (93%)	0.42	13 (5%)	26	32	30, 39, 55, 83	5 (2%)
1	B	240/262 (91%)	0.47	12 (5%)	28	34	28, 38, 52, 63	14 (5%)
1	C	240/262 (91%)	0.70	27 (11%)	5	6	20, 47, 57, 72	51 (21%)
1	D	240/262 (91%)	0.82	41 (17%)	1	1	23, 42, 57, 63	46 (19%)
All	All	964/1048 (91%)	0.60	93 (9%)	8	10	20, 42, 56, 83	116 (12%)

The worst 5 of 93 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	1	MET	6.2
1	D	43	ARG	5.1
1	A	258	ARG	4.9
1	A	257	LEU	4.7
1	B	256	LEU	4.5

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.5 Other polymers ⓘ

There are no such residues in this entry.