Full wwPDB X-ray Structure Validation Report ⓘ

Apr 21, 2024 – 08:31 am BST

PDB ID	:	5FO4
Title	:	Crystal structure of the P.falciparum cytosolic leucyl-tRNA synthetase editing domain (space group P1)
Authors	:	Palencia, A.; Sonoiki, E.; Guo, D.; Ahyong, V.; Dong, C.; Li, X.; Hernandez, V.S.; Gut, J.; Legac, J.; Cooper, R.; Alley, M.R.K.; Freund, Y.R.; DeRisi, J.; Cusack, S.; Rosenthal, P.J.
Deposited on	:	2015-11-18
Resolution	:	1.85 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *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:

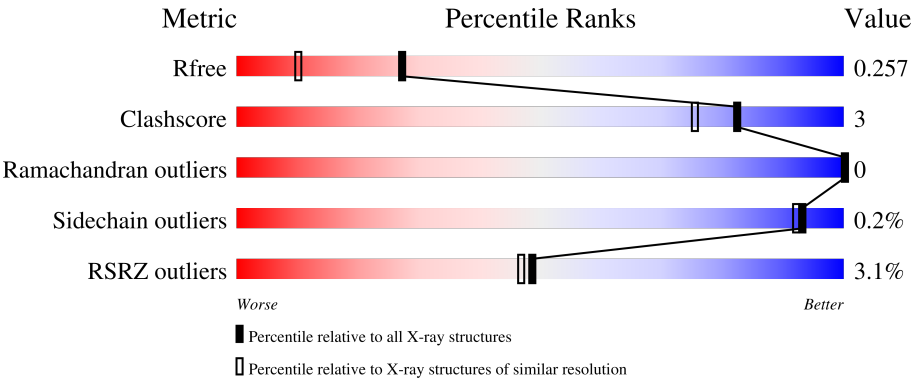
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	374	2% 78% · 19%
1	B	374	3% 82% · 15%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5387 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 LEUCYL TRNA SYNTHASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	302	Total	C	N	O	S	0	4	0
			2508	1620	393	480	15			
1	B	317	Total	C	N	O	S	0	4	0
			2623	1689	413	505	16			

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	269	GLY	-	expression tag	UNP C6KT64
A	270	ALA	-	expression tag	UNP C6KT64
A	271	MET	-	expression tag	UNP C6KT64
A	273	SER	CYS	engineered mutation	UNP C6KT64
A	?	-	ILE	deletion	UNP C6KT64
A	?	-	LYS	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	VAL	deletion	UNP C6KT64
A	?	-	GLN	deletion	UNP C6KT64
A	?	-	ILE	deletion	UNP C6KT64
A	?	-	PRO	deletion	UNP C6KT64
A	?	-	LYS	deletion	UNP C6KT64
A	?	-	SER	deletion	UNP C6KT64
A	?	-	GLU	deletion	UNP C6KT64
A	?	-	ASP	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	THR	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASP	deletion	UNP C6KT64
A	?	-	ASP	deletion	UNP C6KT64
A	?	-	ASP	deletion	UNP C6KT64
A	?	-	THR	deletion	UNP C6KT64
A	?	-	LEU	deletion	UNP C6KT64
A	?	-	GLU	deletion	UNP C6KT64
A	?	-	LYS	deletion	UNP C6KT64

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	LYS	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASP	deletion	UNP C6KT64
A	?	-	VAL	deletion	UNP C6KT64
A	?	-	ILE	deletion	UNP C6KT64
A	?	-	THR	deletion	UNP C6KT64
A	?	-	LYS	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	THR	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	GLU	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	MET	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	ASN	deletion	UNP C6KT64
A	?	-	VAL	deletion	UNP C6KT64
A	520	GLY	TYR	engineered mutation	UNP C6KT64
B	269	GLY	-	expression tag	UNP C6KT64
B	270	ALA	-	expression tag	UNP C6KT64
B	271	MET	-	expression tag	UNP C6KT64
B	273	SER	CYS	engineered mutation	UNP C6KT64
B	?	-	ILE	deletion	UNP C6KT64
B	?	-	LYS	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	VAL	deletion	UNP C6KT64
B	?	-	GLN	deletion	UNP C6KT64
B	?	-	ILE	deletion	UNP C6KT64
B	?	-	PRO	deletion	UNP C6KT64
B	?	-	LYS	deletion	UNP C6KT64
B	?	-	SER	deletion	UNP C6KT64
B	?	-	GLU	deletion	UNP C6KT64
B	?	-	ASP	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	THR	deletion	UNP C6KT64

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASP	deletion	UNP C6KT64
B	?	-	ASP	deletion	UNP C6KT64
B	?	-	ASP	deletion	UNP C6KT64
B	?	-	THR	deletion	UNP C6KT64
B	?	-	LEU	deletion	UNP C6KT64
B	?	-	GLU	deletion	UNP C6KT64
B	?	-	LYS	deletion	UNP C6KT64
B	?	-	LYS	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASP	deletion	UNP C6KT64
B	?	-	VAL	deletion	UNP C6KT64
B	?	-	ILE	deletion	UNP C6KT64
B	?	-	THR	deletion	UNP C6KT64
B	?	-	LYS	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	THR	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	GLU	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	MET	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	ASN	deletion	UNP C6KT64
B	?	-	VAL	deletion	UNP C6KT64
B	520	GLY	TYR	engineered mutation	UNP C6KT64

- Molecule 2 is water.

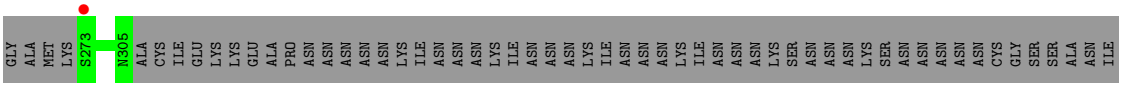
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	138	Total O 138 138	0	0
2	B	118	Total O 118 118	0	0

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 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.

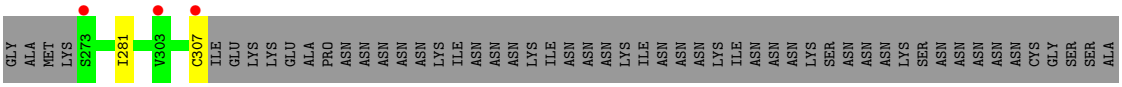
• Molecule 1: LEUCYL TRNA SYNTHASE

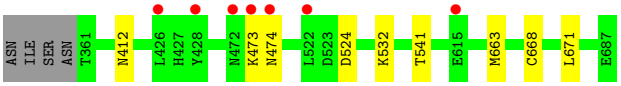
Chain A:



• Molecule 1: LEUCYL TRNA SYNTHASE

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	48.70Å 53.40Å 63.00Å 88.90° 74.30° 89.70°	Depositor
Resolution (Å)	60.64 – 1.85 46.88 – 1.85	Depositor EDS
% Data completeness (in resolution range)	86.9 (60.64-1.85) 86.9 (46.88-1.85)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.37 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.195 , 0.252 0.202 , 0.257	Depositor DCC
R_{free} test set	2151 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å^2)	30.9	Xtriage
Anisotropy	0.409	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 38.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.239 for -h,k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5387	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.62% 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	0.75	0/2568	0.78	1/3452 (0.0%)
1	B	0.77	0/2686	0.79	0/3612
All	All	0.76	0/5254	0.79	1/7064 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	635	ASP	CB-CG-OD1	5.43	123.19	118.30

There are no chirality outliers.

There are no planarity outliers.

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	2508	0	2511	13	0
1	B	2623	0	2613	13	0
2	A	138	0	0	0	0
2	B	118	0	0	0	0
All	All	5387	0	5124	26	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 3.

All (26) 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:663[B]:MET:SD	1:B:668[B]:CYS:SG	2.29	1.30
1:A:663[B]:MET:SD	1:A:668[B]:CYS:SG	2.37	1.22
1:A:663[B]:MET:CE	1:A:668[B]:CYS:SG	2.30	1.19
1:B:663[B]:MET:CE	1:B:668[B]:CYS:SG	2.46	1.03
1:A:663[B]:MET:HE2	1:A:668[B]:CYS:SG	2.08	0.94
1:B:663[B]:MET:HB3	1:B:668[B]:CYS:SG	2.16	0.85
1:B:663[B]:MET:HE2	1:B:668[B]:CYS:SG	2.21	0.81
1:A:663[B]:MET:HB3	1:A:668[B]:CYS:SG	2.20	0.80
1:A:663[B]:MET:HE1	1:A:668[B]:CYS:HA	1.70	0.72
1:A:663[B]:MET:HE1	1:A:668[B]:CYS:SG	2.30	0.69
1:B:663[B]:MET:HE1	1:B:668[B]:CYS:HA	1.75	0.67
1:B:473:LYS:C	1:B:474:ASN:N	2.50	0.65
1:B:663[B]:MET:CB	1:B:668[B]:CYS:SG	2.91	0.58
1:A:663[B]:MET:HE1	1:A:668[B]:CYS:CA	2.34	0.55
1:A:663[B]:MET:CB	1:A:668[B]:CYS:SG	2.95	0.50
1:A:663[B]:MET:CG	1:A:668[B]:CYS:SG	3.01	0.48
1:B:663[B]:MET:CG	1:B:668[B]:CYS:SG	3.02	0.47
1:A:663[B]:MET:CE	1:A:671:LEU:HD12	2.49	0.43
1:A:412:ASN:HA	1:A:532:LYS:HE3	1.99	0.43
1:B:663[B]:MET:HE1	1:B:668[B]:CYS:CA	2.45	0.42
1:B:281:ILE:HG22	1:B:541:THR:HB	2.01	0.41
1:B:473:LYS:O	1:B:524:ASP:HB3	2.20	0.41
1:B:663[B]:MET:CE	1:B:671:LEU:HD12	2.51	0.41
1:B:412:ASN:HA	1:B:532:LYS:HE3	2.02	0.41
1:A:677:ILE:HG12	1:A:684:LEU:HG	2.02	0.40
1:A:590:ASN:O	1:A:594:GLU:HB2	2.21	0.40

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	298/374 (80%)	291 (98%)	7 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	315/374 (84%)	304 (96%)	11 (4%)	0	100 100
All	All	613/748 (82%)	595 (97%)	18 (3%)	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	289/349 (83%)	289 (100%)	0	100 100
1	B	301/349 (86%)	300 (100%)	1 (0%)	92 91
All	All	590/698 (84%)	589 (100%)	1 (0%)	93 92

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	307	CYS

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 (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	473:LYS	C	474:ASN	N	2.50

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, 95^{th} 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	302/374 (80%)	-0.30	9 (2%) 50 48	20, 32, 56, 93	0
1	B	317/374 (84%)	-0.14	10 (3%) 47 45	19, 34, 65, 90	0
All	All	619/748 (82%)	-0.22	19 (3%) 49 47	19, 33, 62, 93	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	463	TYR	5.3
1	B	273	SER	4.7
1	B	522	LEU	4.4
1	B	474	ASN	3.9
1	B	428	TYR	3.7
1	B	615	GLU	3.5
1	A	427	HIS	3.4
1	B	473	LYS	3.4
1	A	426	LEU	3.3
1	A	435	ASN	3.2
1	A	434	VAL	3.0
1	B	307	CYS	2.8
1	B	303	VAL	2.6
1	A	273	SER	2.5
1	A	615	GLU	2.3
1	B	426	LEU	2.3
1	B	472	ASN	2.2
1	A	425	ASN	2.1
1	A	361	THR	2.0

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 (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.