Full wwPDB X-ray Structure Validation Report ⓘ

Apr 21, 2024 – 12:57 pm BST

PDB ID	:	5FOC
Title	:	Crystal structure of the P.falciparum cytosolic leucyl-tRNA synthetase editing domain (space group P21)
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	:	2015-11-19
Resolution	:	1.50 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:

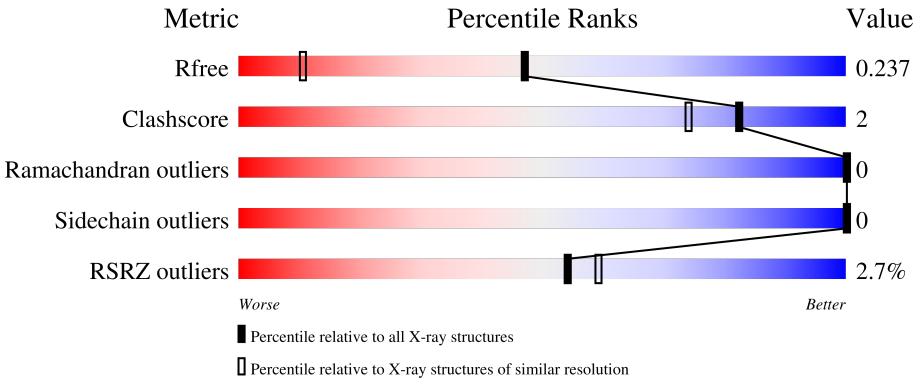
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

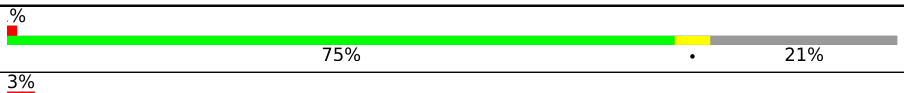
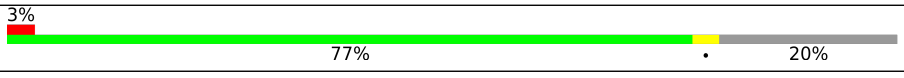
The reported resolution of this entry is 1.50 Å.

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	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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	
1	B	374	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5369 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 SYNTHETASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	297	Total	C	N	O	S	0	7	0
			2486	1608	387	476	15			
1	B	300	Total	C	N	O	S	0	6	0
			2501	1616	390	481	14			

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.

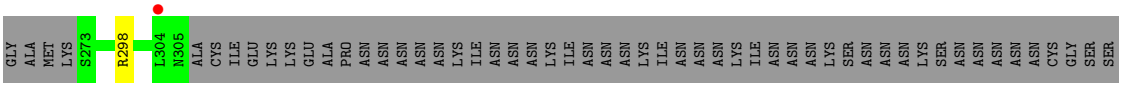
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	219	Total O 219 219	0	0
2	B	163	Total O 163 163	0	0

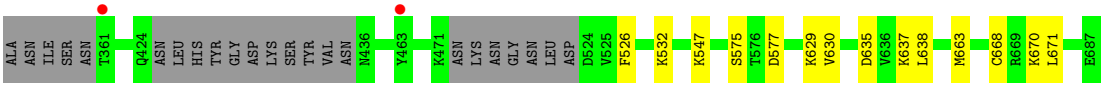
3 Residue-property plots (i)

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 SYNTHETASE

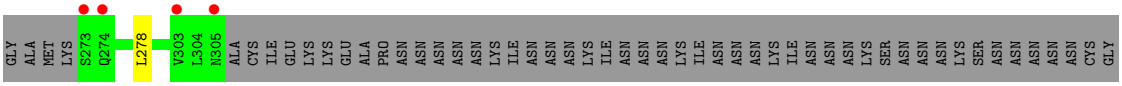
Chain A:

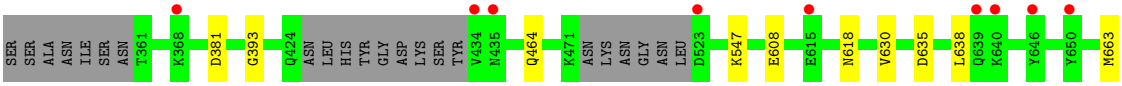




• Molecule 1: LEUCYL-TRNA SYNTHETASE

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	48.60Å 53.49Å 122.79Å 90.00° 97.31° 90.00°	Depositor
Resolution (Å)	50.01 – 1.50 48.97 – 1.45	Depositor EDS
% Data completeness (in resolution range)	99.3 (50.01-1.50) 99.2 (48.97-1.45)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.14 (at 1.45Å)	Xtriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.183 , 0.234 0.192 , 0.237	Depositor DCC
R_{free} test set	5506 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	23.2	Xtriage
Anisotropy	0.257	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 42.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	5369	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 60.83 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.4350e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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	0.78	0/2551	0.86	3/3428 (0.1%)
1	B	0.71	0/2566	0.77	2/3450 (0.1%)
All	All	0.74	0/5117	0.82	5/6878 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	381	ASP	CB-CG-OD2	-6.46	112.48	118.30
1	A	526	PHE	CB-CG-CD2	-5.31	117.09	120.80
1	A	526	PHE	CB-CG-CD1	5.29	124.50	120.80
1	A	298	ARG	NE-CZ-NH2	-5.05	117.78	120.30
1	B	381	ASP	CB-CG-OD1	5.02	122.81	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	2486	0	2496	15	0
1	B	2501	0	2507	10	0
2	A	219	0	0	3	0
2	B	163	0	0	0	0
All	All	5369	0	5003	25	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 2.

All (25) 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:635:ASP:HB3	1:B:638[A]:LEU:HD13	1.80	0.63
1:A:663[A]:MET:HE1	1:A:671:LEU:HD12	1.80	0.63
1:A:630:VAL:CG1	1:A:638[A]:LEU:HD11	2.33	0.59
1:A:575[A]:SER:OG	1:A:577:ASP:OD1	2.22	0.58
1:A:663[A]:MET:CE	1:A:671:LEU:HD12	2.34	0.58
1:B:663:MET:SD	1:B:668[B]:CYS:SG	3.04	0.56
1:B:547:LYS:NZ	1:B:608:GLU:OE1	2.40	0.55
1:B:464[B]:GLN:HE22	1:B:686:SER:H	1.57	0.52
1:B:630:VAL:CG2	1:B:638[A]:LEU:HD11	2.41	0.51
1:A:670:LYS:HE3	2:A:2199:HOH:O	2.12	0.48
1:A:630:VAL:HA	1:A:637:LYS:CE	2.44	0.48
1:B:630:VAL:HG22	1:B:638[A]:LEU:HD11	1.96	0.47
1:A:630:VAL:HA	1:A:637:LYS:HE2	1.97	0.46
1:A:629:LYS:O	1:A:637:LYS:NZ	2.48	0.45
1:A:663[B]:MET:HB2	1:A:668[B]:CYS:SG	2.56	0.45
1:A:635:ASP:HB3	1:A:638[A]:LEU:HD13	1.97	0.45
1:A:630:VAL:HG11	1:A:638[A]:LEU:HD11	1.99	0.44
1:A:532:LYS:NZ	2:A:2083:HOH:O	2.50	0.44
1:A:547:LYS:NZ	2:A:2129:HOH:O	2.51	0.44
1:A:663[A]:MET:HE1	1:A:668[A]:CYS:HA	1.99	0.43
1:A:577:ASP:HA	1:A:638[B]:LEU:HD21	2.00	0.43
1:B:630:VAL:CG1	1:B:638[A]:LEU:HD11	2.50	0.41
1:B:278:LEU:HA	1:B:393:GLY:O	2.22	0.40
1:B:663:MET:HE2	1:B:663:MET:HB3	1.96	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	296/374 (79%)	291 (98%)	5 (2%)	0	100	100
1	B	298/374 (80%)	293 (98%)	5 (2%)	0	100	100
All	All	594/748 (79%)	584 (98%)	10 (2%)	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	287/349 (82%)	287 (100%)	0	100	100
1	B	289/349 (83%)	289 (100%)	0	100	100
All	All	576/698 (82%)	576 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	597	ASN

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)

There are no chain breaks in this entry.

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	297/374 (79%)	-0.15	3 (1%) 82 85	18, 26, 46, 63	0
1	B	300/374 (80%)	0.05	13 (4%) 35 39	19, 30, 54, 71	0
All	All	597/748 (79%)	-0.05	16 (2%) 54 59	18, 28, 50, 71	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	434	VAL	7.8
1	B	523	ASP	5.5
1	A	361	THR	4.9
1	B	303	VAL	4.1
1	B	305	ASN	3.9
1	B	435	ASN	3.6
1	B	615	GLU	3.5
1	B	368	LYS	2.8
1	B	639	GLN	2.6
1	A	304	LEU	2.6
1	B	273	SER	2.4
1	B	650	TYR	2.4
1	A	463	TYR	2.3
1	B	646	TYR	2.3
1	B	640	LYS	2.2
1	B	274	GLN	2.1

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.