Full wwPDB X-ray Structure Validation Report (i)

Jan 3, 2024 – 12:33 pm GMT

PDB ID	:	5FON
Title	:	Crystal structure of the Cryptosporidium muris cytosolic leucyl-tRNA synthetase editing domain (apo structure)
Authors	:	Palencia, A.; Liu, R.J.; Lukarska, M.; Gut, J.; Bougdour, A.; Touquet, B.; Wang, E.D.; Alley, M.R.K.; Rosenthal, P.J.; Hakimi, M.A.; Cusack, S.
Deposited on	:	2015-11-24
Resolution	:	2.70 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 (i) 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 (1)) were used in the production of this report:

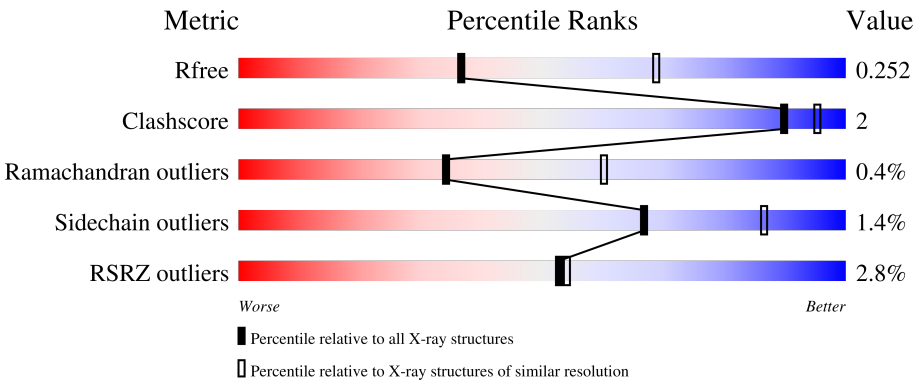
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	291	91% 5% •
1	C	291	3% 83% 7% • 9%
1	D	291	% 86% 7% 7%
2	B	291	6% 84% 8% 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8730 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
			Total	C	N	O	S			
1	A	281	2240	1452	352	421	15	0	0	0
1	C	264	2116	1375	330	396	15	0	0	0
1	D	271	2167	1409	335	408	15	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	251	GLY	-	expression tag	UNP B6AA20
A	252	ALA	-	expression tag	UNP B6AA20
A	253	MET	-	expression tag	UNP B6AA20
C	251	GLY	-	expression tag	UNP B6AA20
C	252	ALA	-	expression tag	UNP B6AA20
C	253	MET	-	expression tag	UNP B6AA20
D	251	GLY	-	expression tag	UNP B6AA20
D	252	ALA	-	expression tag	UNP B6AA20
D	253	MET	-	expression tag	UNP B6AA20

- Molecule 2 is a protein called LEUCYL-TRNA SYNTHETASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	267	2136	1387	333	401	15	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	251	GLY	-	expression tag	UNP B6AA20
B	252	ALA	-	expression tag	UNP B6AA20
B	253	MET	-	expression tag	UNP B6AA20

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	325	GLU	ASP	conflict	UNP B6AA20

- Molecule 3 is water.

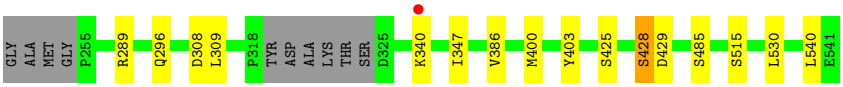
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	19	19	19	0	0
			Total	O		
3	B	14	14	14	0	0
			Total	O		
3	C	15	15	15	0	0
			Total	O		
3	D	23	23	23	0	0

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.

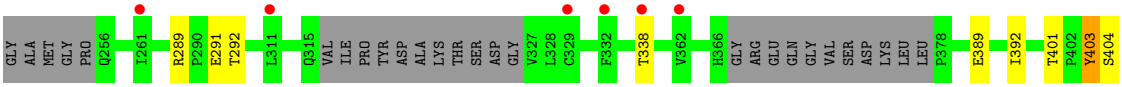
• Molecule 1: LEUCYL-TRNA SYNTHETASE

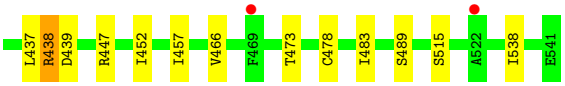
Chain A:



• Molecule 1: LEUCYL-TRNA SYNTHETASE

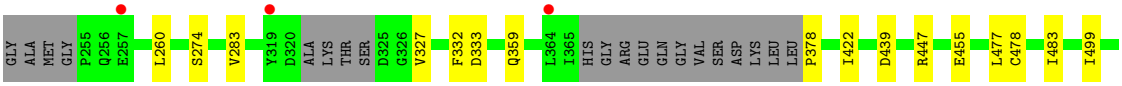
Chain C:





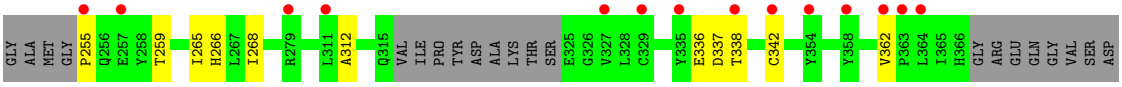
• Molecule 1: LEUCYL-TRNA SYNTHETASE

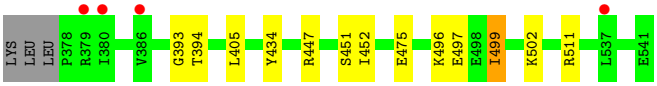
Chain D:



• Molecule 2: LEUCYL-TRNA SYNTHETASE

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	107.71Å 107.71Å 311.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.00 – 2.70 48.17 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.00-2.70) 99.9 (48.17-2.70)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.41 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.203 , 0.256 0.205 , 0.252	Depositor DCC
R_{free} test set	2609 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	78.2	Xtriage
Anisotropy	0.312	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 47.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8730	wwPDB-VP
Average B, all atoms (Å^2)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.02% 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 (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	0.80	0/2286	0.91	2/3087 (0.1%)
1	C	0.78	0/2159	0.91	4/2913 (0.1%)
1	D	0.82	1/2212 (0.0%)	0.91	4/2987 (0.1%)
2	B	0.77	0/2180	0.90	1/2941 (0.0%)
All	All	0.79	1/8837 (0.0%)	0.91	11/11928 (0.1%)

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
2	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	455	GLU	CD-OE1	5.63	1.31	1.25

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	439	ASP	CB-CG-OD1	8.43	125.89	118.30
1	D	439	ASP	CB-CG-OD2	-7.01	112.00	118.30
1	A	289	ARG	NE-CZ-NH2	-6.57	117.01	120.30
2	B	511	ARG	NE-CZ-NH1	5.95	123.28	120.30
1	C	438	ARG	NE-CZ-NH1	5.80	123.20	120.30
1	D	511	ARG	NE-CZ-NH1	5.76	123.18	120.30
1	C	403	TYR	N-CA-C	-5.36	96.52	111.00
1	C	439	ASP	CB-CG-OD1	5.36	123.12	118.30
1	A	403	TYR	C-N-CA	-5.32	108.41	121.70
1	D	447	ARG	NE-CZ-NH1	5.30	122.95	120.30
1	C	289	ARG	NE-CZ-NH2	-5.02	117.79	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	255	PRO	Peptide

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	2240	0	2284	4	0
1	C	2116	0	2156	8	0
1	D	2167	0	2204	8	0
2	B	2136	0	2173	12	1
3	A	19	0	0	0	0
3	B	14	0	0	0	0
3	C	15	0	0	0	0
3	D	23	0	0	1	0
All	All	8730	0	8817	32	1

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 (32) 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:378:PRO:N	3:D:2012:HOH:O	2.21	0.73
1:A:400:MET:SD	1:A:530:LEU:HD21	2.38	0.63
2:B:434:TYR:CZ	2:B:475:GLU:HG3	2.35	0.60
1:D:478:CYS:HA	1:D:483:ILE:HD12	1.84	0.60
2:B:434:TYR:CE1	2:B:475:GLU:HG3	2.41	0.56
1:C:438:ARG:NH2	1:C:483:ILE:O	2.34	0.54
2:B:312:ALA:CB	2:B:362:VAL:HG12	2.42	0.49
2:B:268:ILE:H	2:B:268:ILE:HD12	1.78	0.48
1:C:478:CYS:HA	1:C:483:ILE:HD12	1.95	0.48
2:B:312:ALA:HB2	2:B:362:VAL:HG12	1.95	0.48
1:A:428:SER:HB2	1:A:429:ASP:OD1	2.14	0.48
1:C:437:LEU:HD11	1:C:457:ILE:HG23	1.95	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:265:ILE:HG23	2:B:394:THR:CG2	2.44	0.47
1:D:477:LEU:HD12	1:D:499:ILE:HD12	1.97	0.46
1:A:309:LEU:HD23	1:A:347:ILE:HG12	1.98	0.46
2:B:447:ARG:HA	2:B:452:ILE:HD12	1.97	0.45
1:D:283:VAL:CG1	1:D:422:ILE:HG21	2.47	0.45
1:C:401:THR:HG23	1:C:403:TYR:O	2.17	0.45
2:B:496:LYS:HA	2:B:499:ILE:HG22	2.00	0.44
1:D:260:LEU:HD23	1:D:537:LEU:HD12	2.00	0.44
1:D:327:VAL:HG11	1:D:359:GLN:HA	1.99	0.44
2:B:499:ILE:O	2:B:502:LYS:N	2.50	0.43
1:A:296:GLN:HA	1:A:425:SER:O	2.19	0.43
2:B:393:GLY:O	2:B:394:THR:C	2.57	0.43
1:C:389:GLU:HA	1:C:392:ILE:HD12	2.00	0.42
1:C:447:ARG:HA	1:C:452:ILE:HD12	2.01	0.42
1:D:332:PHE:O	1:D:333:ASP:C	2.59	0.41
1:D:502:LYS:NZ	1:D:506:GLU:OE1	2.54	0.41
1:C:291:GLU:HG2	1:C:292:THR:HG23	2.02	0.41
2:B:266:HIS:CD2	2:B:405:LEU:HD21	2.56	0.41
1:C:466:VAL:HG21	1:C:473:THR:HG21	2.02	0.40
2:B:338:THR:CG2	2:B:342:CYS:SG	3.10	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:336:GLU:OE2	2:B:336:GLU:OE2[7_465]	1.74	0.46

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	277/291 (95%)	263 (95%)	14 (5%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	258/291 (89%)	241 (93%)	17 (7%)	0	100	100
1	D	265/291 (91%)	250 (94%)	14 (5%)	1 (0%)	34	60
2	B	261/291 (90%)	230 (88%)	28 (11%)	3 (1%)	14	34
All	All	1061/1164 (91%)	984 (93%)	73 (7%)	4 (0%)	34	60

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	274	SER
2	B	499	ILE
2	B	451	SER
2	B	337	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	250/256 (98%)	243 (97%)	7 (3%)	43	73
1	C	236/256 (92%)	231 (98%)	5 (2%)	53	80
1	D	242/256 (94%)	242 (100%)	0	100	100
2	B	238/256 (93%)	236 (99%)	2 (1%)	81	93
All	All	966/1024 (94%)	952 (99%)	14 (1%)	67	86

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	308	ASP
1	A	340	LYS
1	A	386	VAL
1	A	428	SER
1	A	485	SER
1	A	515	SER
1	A	540	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	259	THR
2	B	497	GLU
1	C	338	THR
1	C	404	SER
1	C	489	SER
1	C	515	SER
1	C	538	ILE

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

Mol	Chain	Res	Type
2	B	266	HIS
1	D	256	GLN
1	D	315	GLN

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 (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	281/291 (96%)	-0.13	1 (0%)	92	93	54, 76, 121, 149	0
1	C	264/291 (90%)	0.17	8 (3%)	50	51	69, 95, 138, 164	0
1	D	271/291 (93%)	-0.11	3 (1%)	80	82	47, 75, 116, 196	0
2	B	267/291 (91%)	0.28	18 (6%)	17	16	56, 96, 162, 200	0
All	All	1083/1164 (93%)	0.05	30 (2%)	53	54	47, 85, 142, 200	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	537	LEU	5.0
1	D	319	TYR	4.5
2	B	329	CYS	3.9
1	D	257	GLU	3.3
2	B	335	TYR	3.1
2	B	380	ILE	3.1
1	C	332	PHE	3.1
2	B	362	VAL	3.1
2	B	257	GLU	3.0
2	B	364	LEU	3.0
2	B	255	PRO	3.0
1	C	469	PHE	2.8
1	C	362	VAL	2.8
1	D	364	LEU	2.8
2	B	279	ARG	2.6
2	B	327	VAL	2.5
1	C	261	ILE	2.4
2	B	386	VAL	2.4
2	B	363	PRO	2.3
2	B	342	CYS	2.3
1	C	311	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	329	CYS	2.3
2	B	311	LEU	2.2
2	B	338	THR	2.2
2	B	379	ARG	2.1
1	A	340	LYS	2.1
1	C	522	ALA	2.1
2	B	354	TYR	2.1
2	B	358	TYR	2.1
1	C	338	THR	2.1

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 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.