



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 14, 2017 – 03:11 AM EDT

PDB ID : 5U36  
Title : Crystal Structure Of A Mutant M. Jannashii Tyrosyl-tRNA Synthetase  
Authors : Luo, X.; Fu, G.; Zhu, X.; Wilson, I.A.; Wang, F.  
Deposited on : unknown  
Resolution : 3.03 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](mailto: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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20029824
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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-20029824

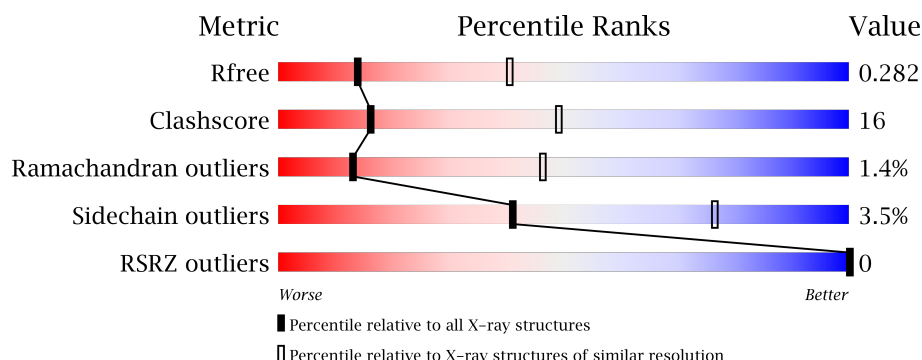
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.03 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))
$R_{free}$	100719	2176 (3.08-3.00)
Clashscore	112137	2542 (3.08-3.00)
Ramachandran outliers	110173	2458 (3.08-3.00)
Sidechain outliers	110143	2461 (3.08-3.00)
RSRZ outliers	101464	2202 (3.08-3.00)

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  $\geq 3$ , 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	314	 69% 23% • 6%
1	B	314	 68% 22% • • 6%

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 4696 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 Tyrosine-tRNA ligase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	294	Total	C	N	O	S	0	0	0
			2348	1506	398	432	12			
1	B	294	Total	C	N	O	S	0	0	0
			2348	1506	398	432	12			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	32	SER	TYR	engineered mutation	UNP Q57834
A	65	ALA	LEU	engineered mutation	UNP Q57834
A	108	LYS	PHE	engineered mutation	UNP Q57834
A	109	HIS	GLN	engineered mutation	UNP Q57834
A	158	GLY	ASP	engineered mutation	UNP Q57834
A	162	LYS	LEU	engineered mutation	UNP Q57834
A	307	LEU	-	expression tag	UNP Q57834
A	308	GLU	-	expression tag	UNP Q57834
A	309	HIS	-	expression tag	UNP Q57834
A	310	HIS	-	expression tag	UNP Q57834
A	311	HIS	-	expression tag	UNP Q57834
A	312	HIS	-	expression tag	UNP Q57834
A	313	HIS	-	expression tag	UNP Q57834
A	314	HIS	-	expression tag	UNP Q57834
B	32	SER	TYR	engineered mutation	UNP Q57834
B	65	ALA	LEU	engineered mutation	UNP Q57834
B	108	LYS	PHE	engineered mutation	UNP Q57834
B	109	HIS	GLN	engineered mutation	UNP Q57834
B	158	GLY	ASP	engineered mutation	UNP Q57834
B	162	LYS	LEU	engineered mutation	UNP Q57834
B	307	LEU	-	expression tag	UNP Q57834
B	308	GLU	-	expression tag	UNP Q57834
B	309	HIS	-	expression tag	UNP Q57834
B	310	HIS	-	expression tag	UNP Q57834
B	311	HIS	-	expression tag	UNP Q57834

*Continued on next page...*

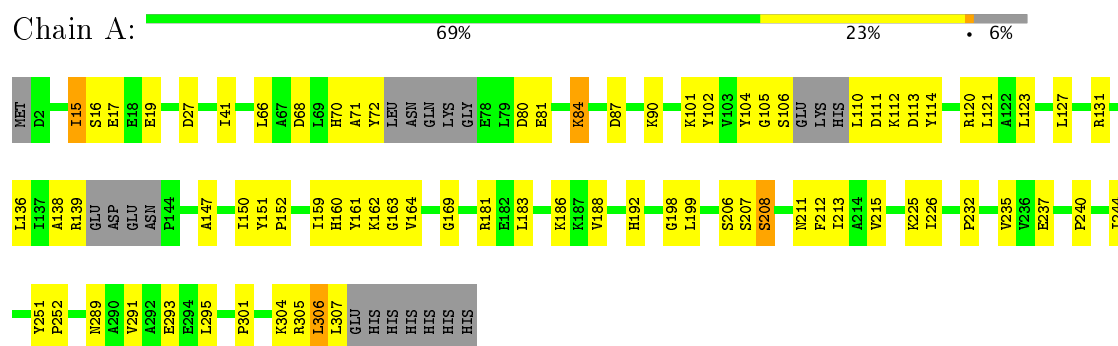
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	312	HIS	-	expression tag	UNP Q57834
B	313	HIS	-	expression tag	UNP Q57834
B	314	HIS	-	expression tag	UNP Q57834

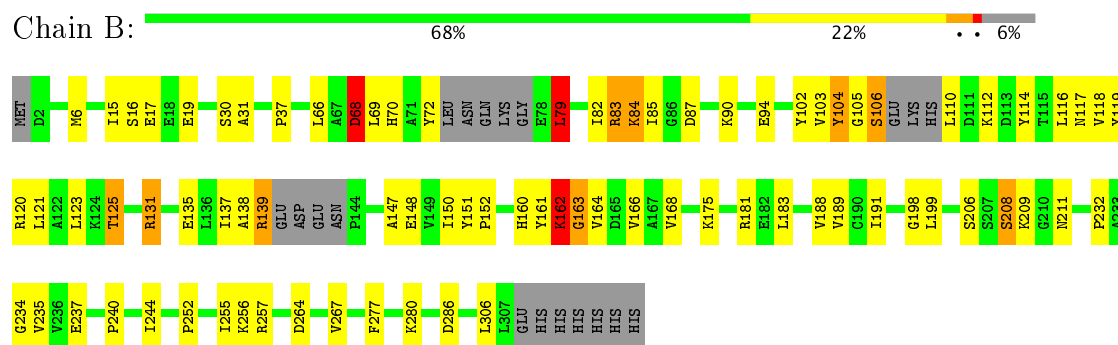
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: Tyrosine-tRNA ligase



#### • Molecule 1: Tyrosine-tRNA ligase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.63Å 130.56Å 59.96Å 90.00° 117.71° 90.00°	Depositor
Resolution (Å)	65.28 – 3.03 41.18 – 3.03	Depositor EDS
% Data completeness (in resolution range)	97.7 (65.28-3.03) 97.7 (41.18-3.03)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.90 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.201 , 0.282 0.201 , 0.282	Depositor DCC
$R_{free}$ test set	759 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	77.9	Xtriage
Anisotropy	0.326	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 35.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.014 for -h-l,k,h 0.014 for l,k,-h-l 0.042 for h,-k,-h-l 0.044 for -h-l,-k,l 0.038 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4696	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.97% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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	0.58	0/2384	0.82	2/3194 (0.1%)
1	B	0.58	0/2384	0.90	9/3194 (0.3%)
All	All	0.58	0/4768	0.86	11/6388 (0.2%)

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	1

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	68	ASP	CB-CA-C	-9.35	91.69	110.40
1	B	83	ARG	C-N-CA	-7.60	102.70	121.70
1	B	68	ASP	CB-CG-OD2	-6.71	112.26	118.30
1	B	79	LEU	CB-CG-CD1	-6.58	99.81	111.00
1	A	306	LEU	CA-CB-CG	6.19	129.53	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	136	LEU	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	2348	0	2439	64	0
1	B	2348	0	2439	90	0
All	All	4696	0	4878	154	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 16.

The worst 5 of 154 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:16:SER:HB2	1:B:19:GLU:CG	1.68	1.22
1:B:79:LEU:CD1	1:B:83:ARG:HE	1.66	1.08
1:B:68:ASP:O	1:B:70:HIS:N	1.88	1.05
1:B:79:LEU:HD11	1:B:83:ARG:HE	0.91	1.04
1:B:16:SER:CB	1:B:19:GLU:CD	2.26	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	286/314 (91%)	272 (95%)	10 (4%)	4 (1%)	13	46
1	B	286/314 (91%)	270 (94%)	12 (4%)	4 (1%)	13	46
All	All	572/628 (91%)	542 (95%)	22 (4%)	8 (1%)	13	46



5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	105	GLY
1	B	105	GLY
1	A	163	GLY
1	A	207	SER
1	B	162	LYS

### 5.3.2 Protein sidechains ⓘ

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	255/274 (93%)	250 (98%)	5 (2%)	60	86
1	B	255/274 (93%)	242 (95%)	13 (5%)	28	65
All	All	510/548 (93%)	492 (96%)	18 (4%)	41	76

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	68	ASP
1	B	79	LEU
1	B	131	ARG
1	B	17	GLU
1	B	30	SER

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

Mol	Chain	Res	Type
1	A	160	HIS

### 5.3.3 RNA ⓘ

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 carbohydrates 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	294/314 (93%)	-0.53	0 100 100	48, 74, 123, 151	0
1	B	294/314 (93%)	-0.51	0 100 100	53, 80, 126, 162	0
All	All	588/628 (93%)	-0.52	0 100 100	48, 77, 126, 162	0

There are no RSRZ outliers to report.

### 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.