



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2018 – 10:42 PM EDT

PDB ID : 5UJI  
Title : Crystal structure of human T2-Tryptophanyl-tRNA synthetase with H130R mutation  
Authors : Xu, X.; Yang, X.-L.  
Deposited on : 2017-01-18  
Resolution : 2.79 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](mailto: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 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 : rb-20031172  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20031172

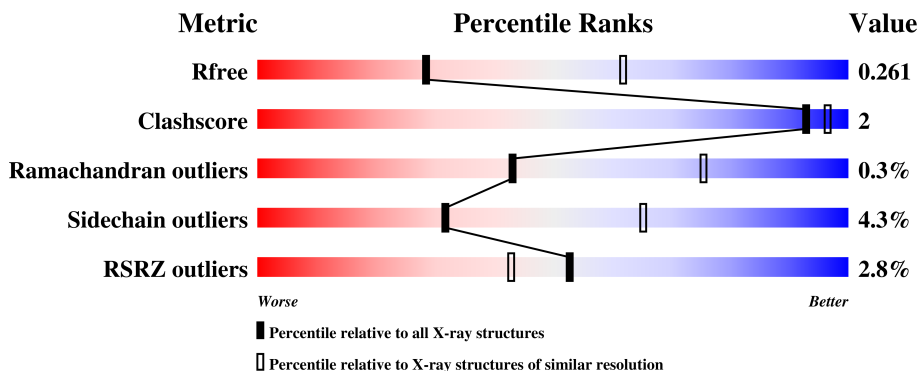
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.79 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))
$R_{free}$	111664	2792 (2.80-2.80)
Clashscore	122126	3209 (2.80-2.80)
Ramachandran outliers	120053	3158 (2.80-2.80)
Sidechain outliers	120020	3160 (2.80-2.80)
RSRZ outliers	108989	2726 (2.80-2.80)

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	388	 2% 85% 6% • 8%
1	B	388	 3% 84% 7% 9%

## 2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 5752 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 Tryptophan-tRNA ligase, cytoplasmic.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	356	2884	1855	486	529	14	0	0	0
1	B	354	2868	1849	483	522	14	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	130	ARG	HIS	engineered mutation	UNP P23381
A	472	LYS	-	expression tag	UNP P23381
A	473	LEU	-	expression tag	UNP P23381
A	474	ALA	-	expression tag	UNP P23381
A	475	ALA	-	expression tag	UNP P23381
A	476	ALA	-	expression tag	UNP P23381
A	477	LEU	-	expression tag	UNP P23381
A	478	GLU	-	expression tag	UNP P23381
A	479	HIS	-	expression tag	UNP P23381
A	480	HIS	-	expression tag	UNP P23381
A	481	HIS	-	expression tag	UNP P23381
A	482	HIS	-	expression tag	UNP P23381
A	483	HIS	-	expression tag	UNP P23381
A	484	HIS	-	expression tag	UNP P23381
B	130	ARG	HIS	engineered mutation	UNP P23381
B	472	LYS	-	expression tag	UNP P23381
B	473	LEU	-	expression tag	UNP P23381
B	474	ALA	-	expression tag	UNP P23381
B	475	ALA	-	expression tag	UNP P23381
B	476	ALA	-	expression tag	UNP P23381
B	477	LEU	-	expression tag	UNP P23381
B	478	GLU	-	expression tag	UNP P23381
B	479	HIS	-	expression tag	UNP P23381
B	480	HIS	-	expression tag	UNP P23381
B	481	HIS	-	expression tag	UNP P23381

*Continued on next page...*

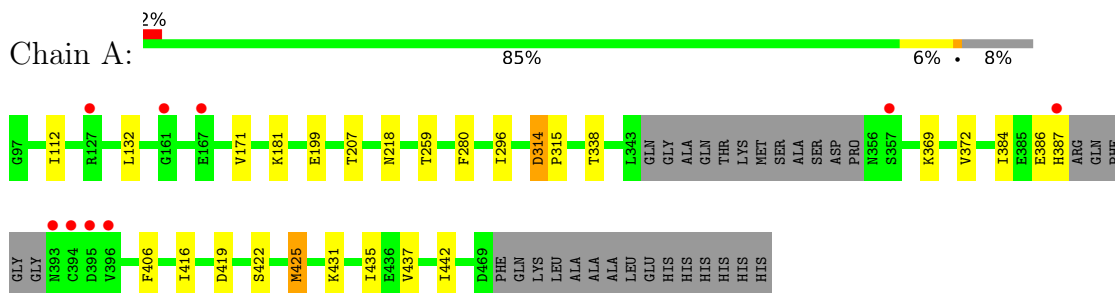
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	482	HIS	-	expression tag	UNP P23381
B	483	HIS	-	expression tag	UNP P23381
B	484	HIS	-	expression tag	UNP P23381

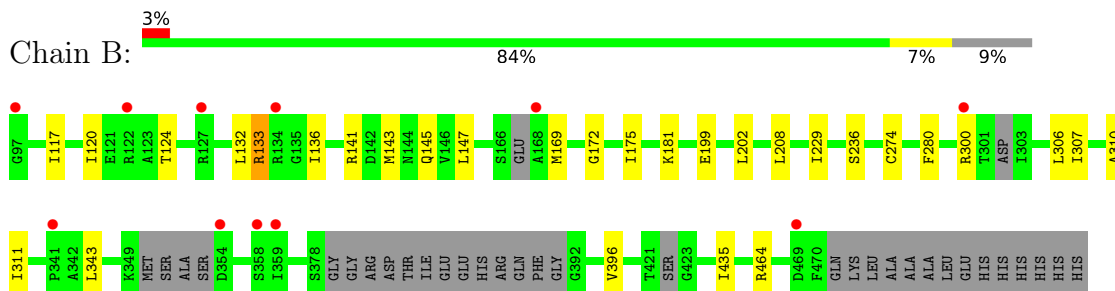
### 3 Residue-property plots [i](#)

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: Tryptophan-tRNA ligase, cytoplasmic



- Molecule 1: Tryptophan-tRNA ligase, cytoplasmic



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	84.89Å 139.06Å 146.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.73 – 2.79 32.73 – 2.79	Depositor EDS
% Data completeness (in resolution range)	98.5 (32.73-2.79) 98.6 (32.73-2.79)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.95 (at 2.81Å)	Xtrriage
Refinement program	REFMAC 5.8.0103	Depositor
R, $R_{free}$	0.219 , 0.264 0.217 , 0.261	Depositor DCC
$R_{free}$ test set	1086 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	59.2	Xtrriage
Anisotropy	0.639	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 22.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.017 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.034 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5752	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.13% 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	0.40	0/2951	0.60	0/3979
1	B	0.40	0/2933	0.62	0/3951
All	All	0.40	0/5884	0.61	0/7930

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2884	0	2852	11	0
1	B	2868	0	2845	8	0
All	All	5752	0	5697	18	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 (18) 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:120:ILE:HD13	1:B:147:LEU:HD21	1.82	0.60
1:B:117:ILE:HG12	1:B:132:LEU:HD13	1.93	0.49

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:133:ARG:HB3	1:B:136:ILE:HD12	1.94	0.49
1:A:171:VAL:HG11	1:A:442:ILE:HD11	1.95	0.49
1:B:172:GLY:HA2	1:B:175:ILE:HD12	1.94	0.48
1:A:425:MET:SD	1:A:425:MET:N	2.88	0.47
1:B:199:GLU:HB2	1:B:280:PHE:CZ	2.49	0.47
1:A:199:GLU:HB2	1:A:280:PHE:CZ	2.50	0.47
1:B:311:ILE:N	1:B:311:ILE:HD12	2.31	0.45
1:A:314:ASP:N	1:A:315:PRO:CD	2.80	0.45
1:A:372:VAL:HB	1:A:435:ILE:HD11	1.99	0.44
1:A:369:LYS:HA	1:A:435:ILE:HD12	2.01	0.43
1:B:307:ILE:HG22	1:B:310:ALA:HB2	2.01	0.43
1:A:259:THR:HA	1:B:274:CYS:HA	2.02	0.42
1:A:431:LYS:O	1:A:435:ILE:HG12	2.20	0.42
1:A:416:ILE:O	1:A:419:ASP:O	2.38	0.41
1:A:296:ILE:HD12	1:A:296:ILE:N	2.36	0.41
1:A:406:PHE:HB3	1:A:437:VAL:CG1	2.51	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	350/388 (90%)	335 (96%)	13 (4%)	2 (1%)	27	60
1	B	342/388 (88%)	323 (94%)	19 (6%)	0	100	100
All	All	692/776 (89%)	658 (95%)	32 (5%)	2 (0%)	43	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	422	SER
1	A	314	ASP



### 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	316/342 (92%)	306 (97%)	10 (3%)	42	76
1	B	314/342 (92%)	297 (95%)	17 (5%)	24	56
All	All	630/684 (92%)	603 (96%)	27 (4%)	32	65

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	112	ILE
1	A	132	LEU
1	A	181	LYS
1	A	207	THR
1	A	218	ASN
1	A	338	THR
1	A	384	ILE
1	A	386	GLU
1	A	387	HIS
1	A	425	MET
1	B	124	THR
1	B	133	ARG
1	B	141	ARG
1	B	143	MET
1	B	145	GLN
1	B	169	MET
1	B	181	LYS
1	B	202	LEU
1	B	208	LEU
1	B	229	ILE
1	B	236	SER
1	B	300	ARG
1	B	306	LEU
1	B	343	LEU
1	B	396	VAL
1	B	435	ILE
1	B	464	ARG

Some 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 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

### 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<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	356/388 (91%)	0.06	9 (2%) 57 47	45, 68, 107, 140	0
1	B	354/388 (91%)	0.08	11 (3%) 49 38	43, 70, 115, 132	0
All	All	710/776 (91%)	0.07	20 (2%) 53 43	43, 69, 111, 140	0

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	394	CYS	5.8
1	A	357	SER	3.9
1	A	393	ASN	3.7
1	B	300	ARG	3.6
1	B	168	ALA	3.1
1	B	122	ARG	3.1
1	A	395	ASP	3.0
1	A	127	ARG	2.7
1	B	97	GLY	2.6
1	B	341	PRO	2.6
1	A	167	GLU	2.5
1	A	396	VAL	2.5
1	B	127	ARG	2.3
1	B	354	ASP	2.3
1	B	358	SER	2.3
1	B	134	ARG	2.2
1	A	161	GLY	2.1
1	B	359	ILE	2.1
1	B	469	ASP	2.1
1	A	387	HIS	2.0

## 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.