wwPDB X-ray Structure Validation Summary Report (i)

Aug 15, 2023 – 09:57 PM EDT

PDB ID	:	1U7X
Title	:	crystal structure of a mutant M. jannashii tyrosyl-tRNA synthetase specific for O-methyl-tyrosine
Authors	:	Zhang, Y.; Wang, L.; Schultz, P.G.; Wilson, I.A.
Deposited on	:	2004-08-04
Resolution	:	3.00 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

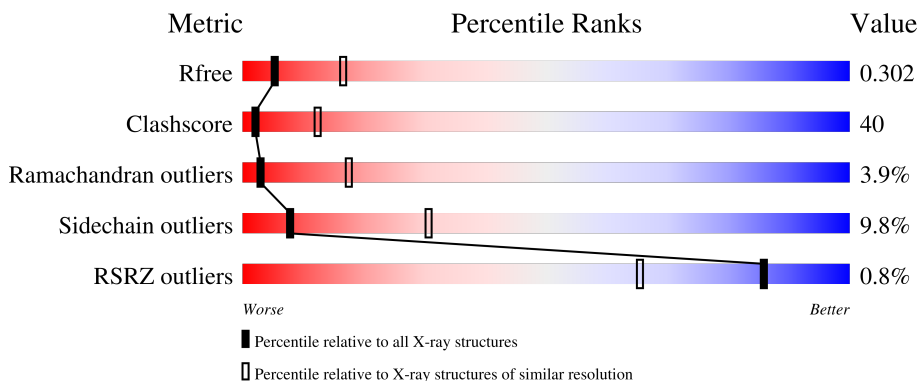|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 3.00 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	Percentile Ranks	Value
Rfree		0.302
Clashscore		40
Ramachandran outliers		3.9%
Sidechain outliers		9.8%
RSRZ outliers		0.8%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	312	2% 38% 52% 7% ··
1	B	312	31% 57% 10% ·

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 4888 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 Tyrosyl-tRNA synthetase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	306	Total	C	N	O	S	0	0	0
			2446	1566	412	455	13			
1	B	306	Total	C	N	O	S	0	0	0
			2441	1563	411	454	13			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	32	GLN	TYR	engineered mutation	UNP Q57834
A	107	THR	GLU	engineered mutation	UNP Q57834
A	158	ALA	ASP	engineered mutation	UNP Q57834
A	162	PRO	LEU	engineered mutation	UNP Q57834
A	307	HIS	-	expression tag	UNP Q57834
A	308	HIS	-	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
B	32	GLN	TYR	engineered mutation	UNP Q57834
B	107	THR	GLU	engineered mutation	UNP Q57834
B	158	ALA	ASP	engineered mutation	UNP Q57834
B	162	PRO	LEU	engineered mutation	UNP Q57834
B	307	HIS	-	expression tag	UNP Q57834
B	308	HIS	-	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
B	312	HIS	-	expression tag	UNP Q57834

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	K 1	0	0

3 Residue-property plots ⓘ

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: Tyrosyl-tRNA synthetase

Chain A:

• Molecule 1: Tyrosyl-tRNA synthetase

Chain B:
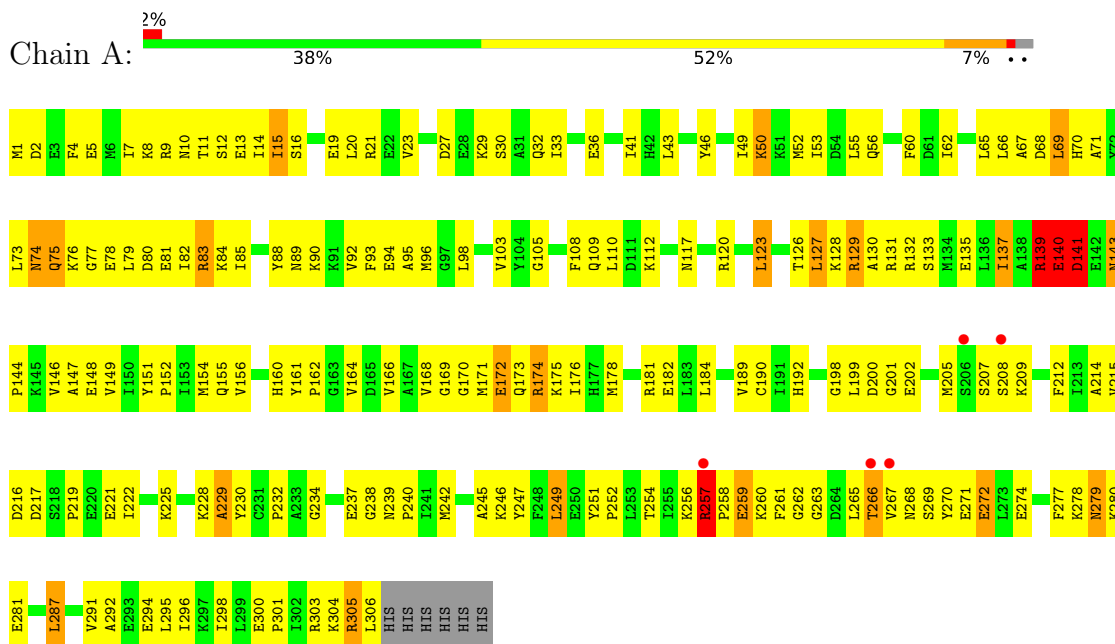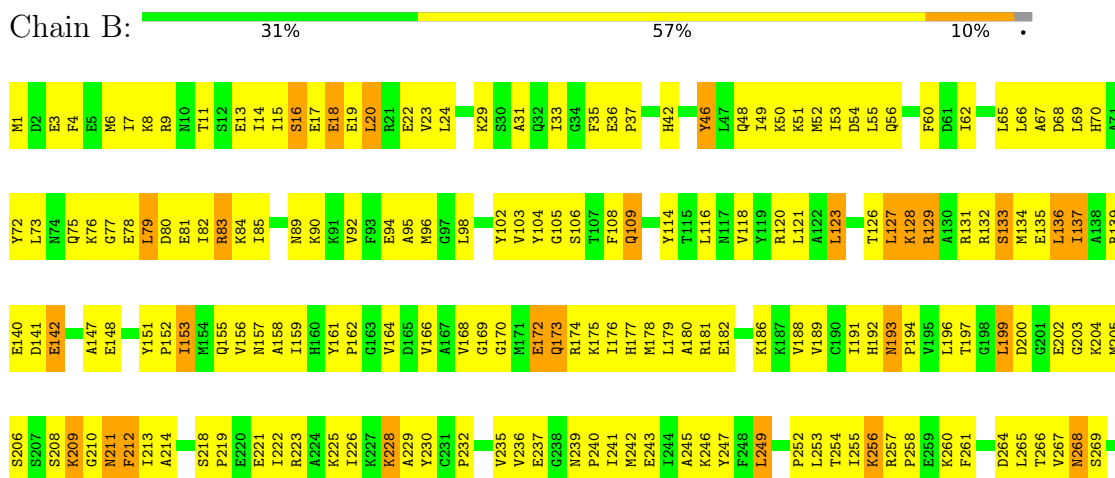
E272 L273 E274 S275 L276 F277 K278 N279 K280 E281 L282 H283 L287 K288 N289 A290 V291 A292 E293 E294 L295 I296 K297 I298 L299 E300 P301 I302 R303 K304 R305 L306 HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	45.37Å 185.00Å 93.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.25 – 3.00 46.25 – 3.00	Depositor EDS
% Data completeness (in resolution range)	89.7 (46.25-3.00) 89.6 (46.25-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.15	Depositor
$< I/\sigma(I) >$ [1]	2.21 (at 3.01Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.229 , 0.308 0.226 , 0.302	Depositor DCC
R_{free} test set	1471 reflections (9.03%)	wwPDB-VP
Wilson B-factor (Å^2)	39.3	Xtriage
Anisotropy	0.621	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 61.2	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	4888	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.55% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: K

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	0.50	1/2487 (0.0%)	0.76	3/3341 (0.1%)
1	B	0.45	0/2482	0.69	1/3336 (0.0%)
All	All	0.47	1/4969 (0.0%)	0.73	4/6677 (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
1	A	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	140	GLU	C-N	9.66	1.56	1.34

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	139	ARG	O-C-N	-15.18	98.41	122.70
1	A	139	ARG	CA-C-N	7.66	134.06	117.20
1	A	140	GLU	C-N-CA	-5.95	106.83	121.70
1	B	210	GLY	N-CA-C	5.86	127.75	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Res	Type	Group
1	A	139	ARG	Mainchain

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	2446	0	2523	190	0
1	B	2441	0	2509	214	0
2	A	1	0	0	0	0
All	All	4888	0	5032	397	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 40.

The worst 5 of 397 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:A:176:ILE:HD12	1:A:176:ILE:H	1.12	1.11
1:A:140:GLU:O	1:A:141:ASP:O	1.70	1.07
1:A:139:ARG:O	1:A:140:GLU:HB3	1.57	1.01
1:A:279:ASN:HD21	1:A:281:GLU:HB2	1.24	1.01
1:A:245:ALA:HA	1:A:249:LEU:HD22	1.41	0.98

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	304/312 (97%)	255 (84%)	35 (12%)	14 (5%)	2	14
1	B	304/312 (97%)	263 (86%)	31 (10%)	10 (3%)	4	21
All	All	608/624 (97%)	518 (85%)	66 (11%)	24 (4%)	3	17

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	137	ILE
1	A	140	GLU
1	A	141	ASP
1	B	79	LEU
1	B	209	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	266/273 (97%)	244 (92%)	22 (8%)	11	39
1	B	264/273 (97%)	234 (89%)	30 (11%)	5	24
All	All	530/546 (97%)	478 (90%)	52 (10%)	8	30

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	127	LEU
1	B	172	GLU
1	B	289	ASN
1	B	128	LYS
1	B	136	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	74	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	155	GLN
1	B	268	ASN
1	B	193	ASN
1	B	211	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 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	306/312 (98%)	-0.17	5 (1%)	72	44	2, 29, 80, 132	0
1	B	306/312 (98%)	-0.14	0	100	100	10, 38, 74, 98	0
All	All	612/624 (98%)	-0.16	5 (0%)	86	65	2, 34, 76, 132	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	208	SER	2.5
1	A	206	SER	2.5
1	A	257	ARG	2.4
1	A	267	VAL	2.4
1	A	266	THR	2.2

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 ⓘ

There are no monosaccharides in this entry.

6.4   Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	K	A	500	1/1	0.96	0.11	6,6,6,6	0

6.5 Other polymers ⓘ

There are no such residues in this entry.