



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 02:17 am GMT

PDB ID : 3WQY  
Title : Crystal structure of Archaeoglobus fulgidus alanyl-tRNA synthetase in complex with wild-type tRNA(Ala) having G3.U70  
Authors : Naganuma, M.; Sekine, S.; Yokoyama, S.  
Deposited on : 2014-02-05  
Resolution : 3.30 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
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
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)	:	recalc28949

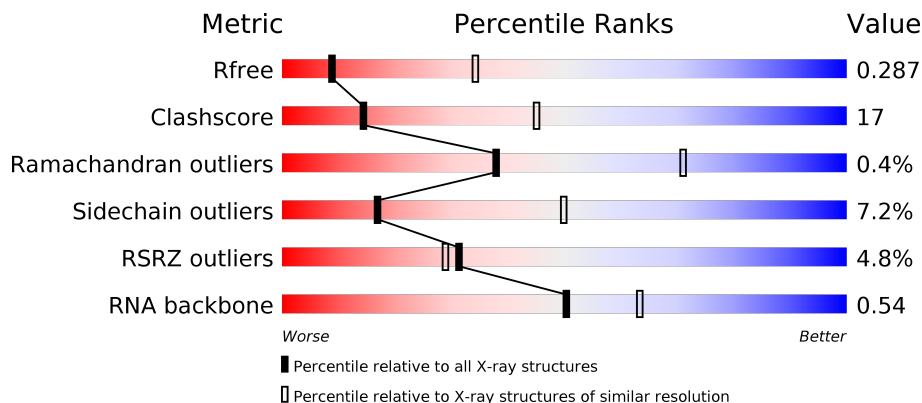
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	1034 (3.36-3.24)
Clashscore	112137	1100 (3.36-3.24)
Ramachandran outliers	110173	1081 (3.36-3.24)
Sidechain outliers	110143	1080 (3.36-3.24)
RSRZ outliers	101464	1039 (3.36-3.24)
RNA backbone	2435	1111 (3.80-2.80)

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	906	<div> <div>5%</div> <div>63%</div> <div>34%</div> <div>.</div> </div>
1	B	906	<div> <div>4%</div> <div>60%</div> <div>36%</div> <div>.</div> </div>
2	C	75	<div> <div>%</div> <div>51%</div> <div>37%</div> <div>12%</div> </div>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 16096 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.

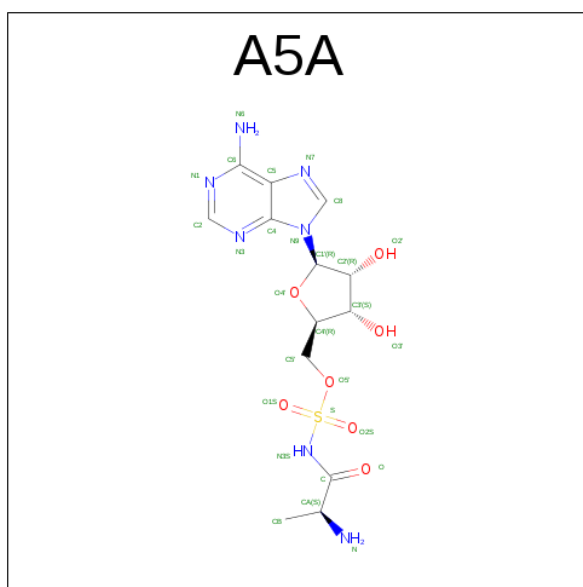
- Molecule 1 is a protein called Alanine-tRNA ligase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	905	Total	C	N	O	S	0	0	0
			7209	4573	1248	1357	31			
1	B	905	Total	C	N	O	S	0	0	0
			7212	4576	1248	1356	32			

- Molecule 2 is a RNA chain called RNA (75-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	75	Total	C	N	O	P	0	0	0
			1609	715	295	524	75			

- Molecule 3 is '5'-O-(N-(L-ALANYL)-SULFAMOYL)ADENOSINE (three-letter code: A5A) (formula: C<sub>13</sub>H<sub>19</sub>N<sub>7</sub>O<sub>7</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			28	13	7	7	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	N	O	S	0	0
			28	13	7	7	1		

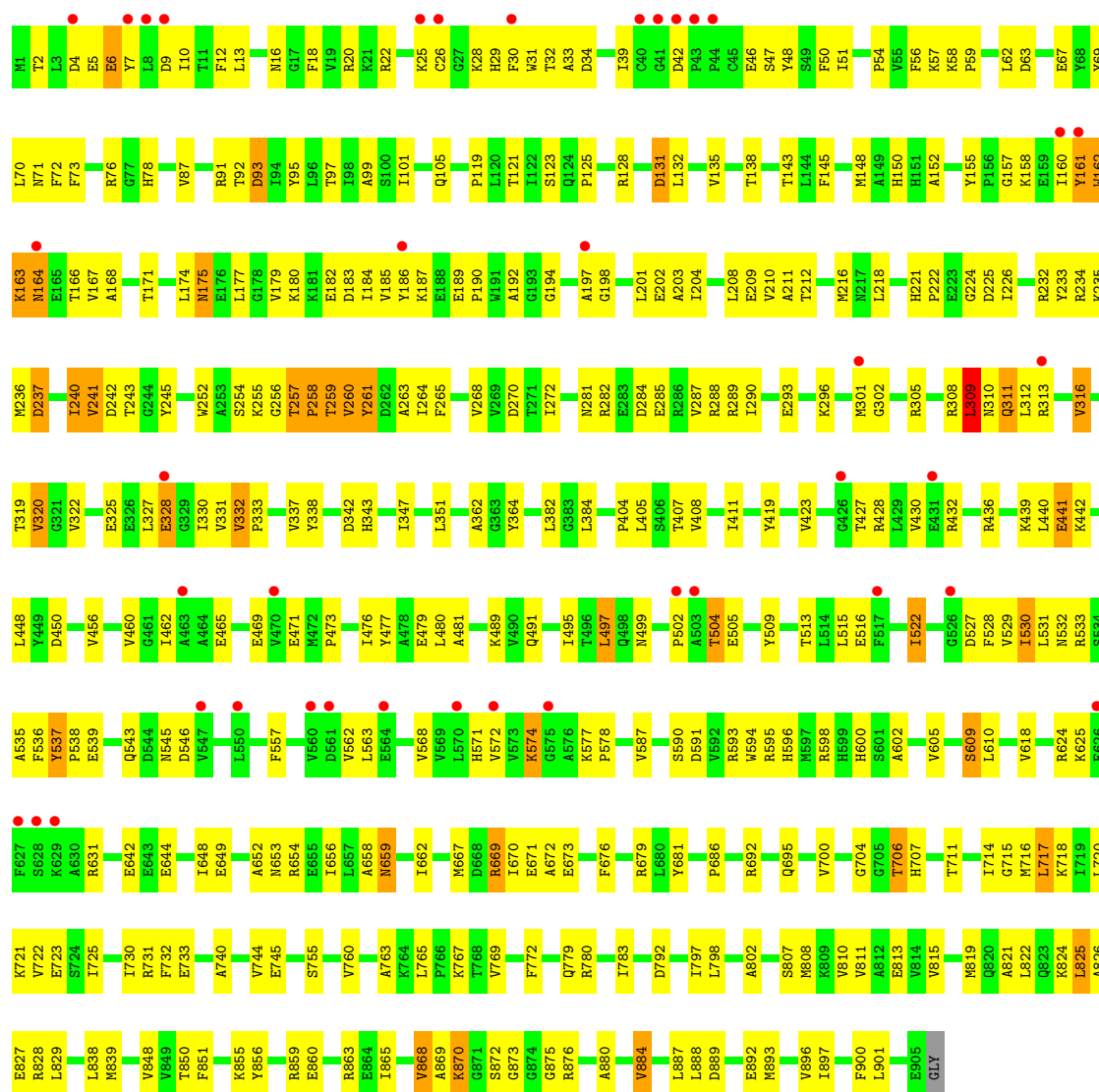
- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	3	Total	Mg	0	0
			3	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	O	0	0
			2	2		
5	B	5	Total	O	0	0
			5	5		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.74Å 169.84Å 175.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.77 – 3.30 41.76 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.5 (41.77-3.30) 99.5 (41.76-3.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.14	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.17 (at 3.32Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, $R_{free}$	0.250 , 0.289 0.251 , 0.287	Depositor DCC
$R_{free}$ test set	2328 reflections (5.14%)	DCC
Wilson B-factor (Å <sup>2</sup> )	75.9	Xtriage
Anisotropy	0.447	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 43.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	0.069 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	16096	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	96.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.66% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: A5A, MG

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	0.27	0/7347	0.47	0/9909
1	B	0.31	0/7350	0.50	1/9914 (0.0%)
2	C	0.26	0/1798	0.83	0/2802
All	All	0.29	0/16495	0.54	1/22625 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	309	LEU	CA-CB-CG	5.77	128.56	115.30

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	7209	0	7201	250	0
1	B	7212	0	7210	267	0
2	C	1609	0	816	34	0
3	A	28	0	19	5	0
3	B	28	0	19	0	0
4	C	3	0	0	0	0
5	A	2	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	5	0	0	0	0
All	All	16096	0	15265	542	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 17.

The worst 5 of 542 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)
2:C:1636:C:H1'	2:C:1637:G:OP2	1.47	1.11
1:B:10:ILE:HG22	1:B:164:ASN:HD22	1.12	1.06
1:B:18:PHE:HB3	1:B:31:TRP:HE1	1.27	0.96
2:C:1637:G:H2'	2:C:1638:A:O4'	1.67	0.95
1:B:20:ARG:HE	1:B:31:TRP:HZ3	1.11	0.94

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	903/906 (100%)	864 (96%)	35 (4%)	4 (0%)	38	71
1	B	903/906 (100%)	855 (95%)	44 (5%)	4 (0%)	38	71
All	All	1806/1812 (100%)	1719 (95%)	79 (4%)	8 (0%)	38	71

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	44	PRO
1	A	544	ASP
1	B	502	PRO

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	870	LYS
1	B	260	VAL

### 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	766/767 (100%)	716 (94%)	50 (6%)	20	55
1	B	767/767 (100%)	707 (92%)	60 (8%)	15	46
All	All	1533/1534 (100%)	1423 (93%)	110 (7%)	17	50

5 of 110 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	862	LEU
1	B	162	TRP
1	B	722	VAL
1	B	42	ASP
1	B	93	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	29	HIS
1	B	105	GLN
1	B	596	HIS
1	B	16	ASN
1	B	571	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	74/75 (98%)	20 (27%)	0

5 of 20 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	C	1603	G
2	C	1609	A
2	C	1613	C
2	C	1616	C
2	C	1619	G

There are no RNA pucker outliers to report.

## 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](#)

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
3	A5A	A	1001	-	25,30,30	2.76	9 (36%)	29,45,45	3.39	7 (24%)
3	A5A	B	1001	-	25,30,30	2.84	9 (36%)	29,45,45	3.13	8 (27%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	A5A	A	1001	-	-	0/14/35/35	0/3/3/3

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	A5A	B	1001	-	-	0/14/35/35	0/3/3/3

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1001	A5A	O5'-S	-8.75	1.46	1.59
3	A	1001	A5A	O5'-S	-7.88	1.48	1.59
3	B	1001	A5A	C2'-C1'	-4.36	1.46	1.53
3	A	1001	A5A	C2'-C1'	-3.99	1.47	1.53
3	B	1001	A5A	C2'-C3'	-3.45	1.44	1.53

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1001	A5A	O2S-S-O1S	-12.69	108.33	121.30
3	B	1001	A5A	O2S-S-O1S	-10.62	110.44	121.30
3	B	1001	A5A	N3-C2-N1	-10.35	119.85	128.86
3	A	1001	A5A	N3-C2-N1	-9.36	120.70	128.86
3	A	1001	A5A	CB-CA-C	-3.22	103.97	110.15

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1001	A5A	5	0

## 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	905/906 (99%)	0.13	49 (5%) 26 25	22, 93, 184, 291	0
1	B	905/906 (99%)	0.20	40 (4%) 35 33	18, 99, 170, 282	0
2	C	75/75 (100%)	-0.48	1 (1%) 77 75	27, 51, 129, 240	0
All	All	1885/1887 (99%)	0.14	90 (4%) 31 29	18, 95, 179, 291	0

The worst 5 of 90 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	628	SER	8.6
1	B	503	ALA	8.0
1	B	42	ASP	8.0
1	A	487	ALA	6.6
1	A	488	GLU	6.6

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	A5A	B	1001	28/28	0.93	0.28	0.49	46,70,102,105	0
3	A5A	A	1001	28/28	0.95	0.27	-0.01	51,66,91,96	0
4	MG	C	1703	1/1	0.90	0.14	-	63,63,63,63	0
4	MG	C	1702	1/1	0.97	0.13	-	13,13,13,13	0
4	MG	C	1701	1/1	0.60	0.62	-	52,52,52,52	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.