



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 13, 2018 – 01:27 pm GMT

PDB ID : 1YFS  
Title : The crystal structure of alanyl-tRNA synthetase in complex with L-alanine  
Authors : Swairjo, M.A.; Schimmel, P.R.  
Deposited on : 2005-01-03  
Resolution : 2.08 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

<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  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk31020  
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) : trunk31020

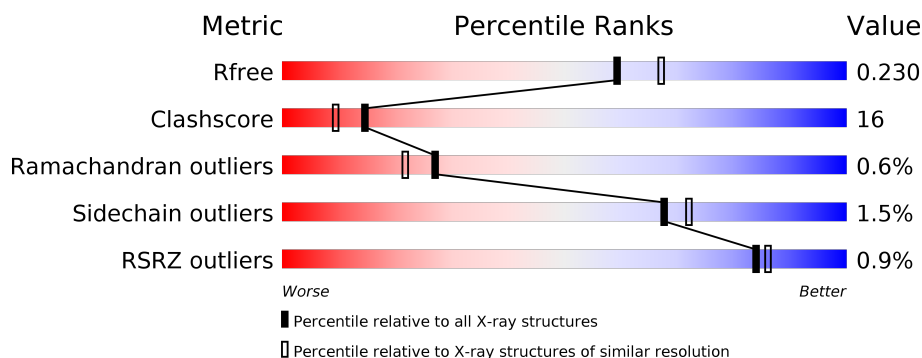
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.08 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	5409 (2.10-2.06)
Clashscore	122126	5943 (2.10-2.06)
Ramachandran outliers	120053	5886 (2.10-2.06)
Sidechain outliers	120020	5887 (2.10-2.06)
RSRZ outliers	108989	5279 (2.10-2.06)

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	465	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: -10px; left: 0;">%</div> <div style="position: absolute; top: 0; left: 0; width: 100%; height: 100%; background: linear-gradient(to right, red 0%, orange 10%, yellow 29%, green 66%, grey 100%);"></div> <div style="position: absolute; bottom: -10px; left: 0;">66% 29% . .</div> </div> </div>
1	B	465	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: 0; left: 0; width: 100%; height: 100%; background: linear-gradient(to right, green 0%, yellow 31%, grey 100%);"></div> <div style="position: absolute; bottom: -10px; left: 0;">64% 31% . .</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7654 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 Alanyl-tRNA synthetase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	448	Total	C	N	O	S	0	0	0
			3641	2339	615	675	12			
1	B	448	Total	C	N	O	S	0	0	0
			3641	2339	615	675	12			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	454	ALA	-	CLONING ARTIFACT	UNP O67323
A	455	ALA	-	CLONING ARTIFACT	UNP O67323
A	456	ALA	-	CLONING ARTIFACT	UNP O67323
A	457	LEU	-	CLONING ARTIFACT	UNP O67323
A	458	GLU	-	CLONING ARTIFACT	UNP O67323
A	459	HIS	-	EXPRESSION TAG	UNP O67323
A	460	HIS	-	EXPRESSION TAG	UNP O67323
A	461	HIS	-	EXPRESSION TAG	UNP O67323
A	462	HIS	-	EXPRESSION TAG	UNP O67323
A	463	HIS	-	EXPRESSION TAG	UNP O67323
A	464	HIS	-	EXPRESSION TAG	UNP O67323
B	454	ALA	-	CLONING ARTIFACT	UNP O67323
B	455	ALA	-	CLONING ARTIFACT	UNP O67323
B	456	ALA	-	CLONING ARTIFACT	UNP O67323
B	457	LEU	-	CLONING ARTIFACT	UNP O67323
B	458	GLU	-	CLONING ARTIFACT	UNP O67323
B	459	HIS	-	EXPRESSION TAG	UNP O67323
B	460	HIS	-	EXPRESSION TAG	UNP O67323
B	461	HIS	-	EXPRESSION TAG	UNP O67323
B	462	HIS	-	EXPRESSION TAG	UNP O67323
B	463	HIS	-	EXPRESSION TAG	UNP O67323
B	464	HIS	-	EXPRESSION TAG	UNP O67323

- Molecule 2 is ALANINE (three-letter code: ALA) (formula: C<sub>3</sub>H<sub>7</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			6	3	1	2		
2	B	1	Total	C	N	O	0	0
			6	3	1	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	183	Total	O	0	0
			183	183		
3	B	177	Total	O	0	0
			177	177		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	173.03Å 73.87Å 74.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.08 36.93 – 2.08	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.00-2.08) 99.4 (36.93-2.08)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	7.97 (at 2.08Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.189 , 0.236 0.185 , 0.230	Depositor DCC
$R_{free}$ test set	5804 reflections (10.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.2	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.21 , 23.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.20$	Xtriage
Estimated twinning fraction	0.306 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7654	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% 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

### 5.1 Standard geometry

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	0.34	0/3725	0.55	0/5016
1	B	0.33	0/3725	0.53	0/5016
All	All	0.34	0/7450	0.54	0/10032

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

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	3641	0	3601	116	0
1	B	3641	0	3601	122	0
2	A	6	0	4	0	0
2	B	6	0	4	0	0
3	A	183	0	0	6	0
3	B	177	0	0	7	0
All	All	7654	0	7210	238	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 238 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:326:ILE:HG22	1:A:327:MET:HE3	1.39	0.98
1:B:26:ALA:H	1:B:65:GLN:HE22	1.11	0.97
1:B:300:ARG:HA	1:B:303:MET:HE3	1.49	0.93
1:A:26:ALA:H	1:A:65:GLN:HE22	1.20	0.86
1:A:421:ARG:HD3	1:A:445:LEU:HD11	1.64	0.78

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	444/465 (96%)	420 (95%)	22 (5%)	2 (0%)	31	27
1	B	444/465 (96%)	416 (94%)	25 (6%)	3 (1%)	24	18
All	All	888/930 (96%)	836 (94%)	47 (5%)	5 (1%)	27	22

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	LEU
1	A	198	MET
1	B	410	LEU
1	B	2	LEU
1	B	78	GLU

### 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	388/404 (96%)	383 (99%)	5 (1%)	71	76
1	B	388/404 (96%)	381 (98%)	7 (2%)	62	66
All	All	776/808 (96%)	764 (98%)	12 (2%)	67	72

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	43	MET
1	B	69	ARG
1	B	371	LEU
1	A	427	HIS
1	B	165	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	420	GLN
1	B	65	GLN
1	B	162	GLN
1	A	368	GLN
1	B	215	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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

2 ligands are modelled in this entry.

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	$\# Z  > 2$
2	ALA	A	1500	-	2,5,5	0.44	0	2,6,6	0.55	0
2	ALA	B	2500	-	2,5,5	0.57	0	2,6,6	0.56	0

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
2	ALA	A	1500	-	-	0/0/4/4	0/0/0/0
2	ALA	B	2500	-	-	0/0/4/4	0/0/0/0

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 [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	448/465 (96%)	-0.48	6 (1%) 77 80	7, 18, 34, 54	0
1	B	448/465 (96%)	-0.49	2 (0%) 92 93	9, 20, 33, 50	0
All	All	896/930 (96%)	-0.48	8 (0%) 84 86	7, 19, 34, 54	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	427	HIS	3.0
1	B	69	ARG	3.0
1	B	427	HIS	2.5
1	A	425	ARG	2.5
1	A	424	ALA	2.2

### 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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
2	ALA	A	1500	6/6	0.96	0.11	20,22,24,24	0
2	ALA	B	2500	6/6	0.97	0.09	16,18,19,22	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.