



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:07 AM GMT

PDB ID : 2ZNJ  
Title : Crystal structure of Pyrrolysyl-tRNA synthetase from *Desulfitobacterium hafniense*  
Authors : Nozawa, K.; Araiso, Y.; Soll, D.; Ishitani, R.; Nureki, O.  
Deposited on : 2008-04-25  
Resolution : 2.50 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 (RC4), CSD as536be (2015)  
Xtrriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

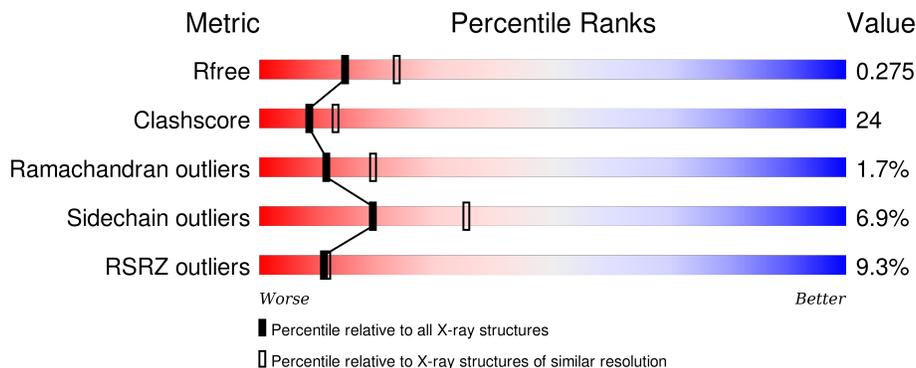
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
$R_{free}$	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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	308	 10% 56% 31% • 10%
1	B	308	 11% 57% 28% • • 9%
1	C	308	 4% 58% 30% • 10%

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 7189 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 Putative uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	278	2255	1434	396	413	12	0	0	0
1	B	279	2263	1440	397	414	12	0	0	0
1	C	278	2255	1434	396	413	12	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP B0S4P3
A	-18	GLY	-	EXPRESSION TAG	UNP B0S4P3
A	-17	SER	-	EXPRESSION TAG	UNP B0S4P3
A	-16	SER	-	EXPRESSION TAG	UNP B0S4P3
A	-15	HIS	-	EXPRESSION TAG	UNP B0S4P3
A	-14	HIS	-	EXPRESSION TAG	UNP B0S4P3
A	-13	HIS	-	EXPRESSION TAG	UNP B0S4P3
A	-12	HIS	-	EXPRESSION TAG	UNP B0S4P3
A	-11	HIS	-	EXPRESSION TAG	UNP B0S4P3
A	-10	HIS	-	EXPRESSION TAG	UNP B0S4P3
A	-9	SER	-	EXPRESSION TAG	UNP B0S4P3
A	-8	SER	-	EXPRESSION TAG	UNP B0S4P3
A	-7	GLY	-	EXPRESSION TAG	UNP B0S4P3
A	-6	LEU	-	EXPRESSION TAG	UNP B0S4P3
A	-5	VAL	-	EXPRESSION TAG	UNP B0S4P3
A	-4	PRO	-	EXPRESSION TAG	UNP B0S4P3
A	-3	ARG	-	EXPRESSION TAG	UNP B0S4P3
A	-2	GLY	-	EXPRESSION TAG	UNP B0S4P3
A	-1	SER	-	EXPRESSION TAG	UNP B0S4P3
A	0	HIS	-	EXPRESSION TAG	UNP B0S4P3
B	-19	MET	-	EXPRESSION TAG	UNP B0S4P3
B	-18	GLY	-	EXPRESSION TAG	UNP B0S4P3
B	-17	SER	-	EXPRESSION TAG	UNP B0S4P3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	SER	-	EXPRESSION TAG	UNP B0S4P3
B	-15	HIS	-	EXPRESSION TAG	UNP B0S4P3
B	-14	HIS	-	EXPRESSION TAG	UNP B0S4P3
B	-13	HIS	-	EXPRESSION TAG	UNP B0S4P3
B	-12	HIS	-	EXPRESSION TAG	UNP B0S4P3
B	-11	HIS	-	EXPRESSION TAG	UNP B0S4P3
B	-10	HIS	-	EXPRESSION TAG	UNP B0S4P3
B	-9	SER	-	EXPRESSION TAG	UNP B0S4P3
B	-8	SER	-	EXPRESSION TAG	UNP B0S4P3
B	-7	GLY	-	EXPRESSION TAG	UNP B0S4P3
B	-6	LEU	-	EXPRESSION TAG	UNP B0S4P3
B	-5	VAL	-	EXPRESSION TAG	UNP B0S4P3
B	-4	PRO	-	EXPRESSION TAG	UNP B0S4P3
B	-3	ARG	-	EXPRESSION TAG	UNP B0S4P3
B	-2	GLY	-	EXPRESSION TAG	UNP B0S4P3
B	-1	SER	-	EXPRESSION TAG	UNP B0S4P3
B	0	HIS	-	EXPRESSION TAG	UNP B0S4P3
C	-19	MET	-	EXPRESSION TAG	UNP B0S4P3
C	-18	GLY	-	EXPRESSION TAG	UNP B0S4P3
C	-17	SER	-	EXPRESSION TAG	UNP B0S4P3
C	-16	SER	-	EXPRESSION TAG	UNP B0S4P3
C	-15	HIS	-	EXPRESSION TAG	UNP B0S4P3
C	-14	HIS	-	EXPRESSION TAG	UNP B0S4P3
C	-13	HIS	-	EXPRESSION TAG	UNP B0S4P3
C	-12	HIS	-	EXPRESSION TAG	UNP B0S4P3
C	-11	HIS	-	EXPRESSION TAG	UNP B0S4P3
C	-10	HIS	-	EXPRESSION TAG	UNP B0S4P3
C	-9	SER	-	EXPRESSION TAG	UNP B0S4P3
C	-8	SER	-	EXPRESSION TAG	UNP B0S4P3
C	-7	GLY	-	EXPRESSION TAG	UNP B0S4P3
C	-6	LEU	-	EXPRESSION TAG	UNP B0S4P3
C	-5	VAL	-	EXPRESSION TAG	UNP B0S4P3
C	-4	PRO	-	EXPRESSION TAG	UNP B0S4P3
C	-3	ARG	-	EXPRESSION TAG	UNP B0S4P3
C	-2	GLY	-	EXPRESSION TAG	UNP B0S4P3
C	-1	SER	-	EXPRESSION TAG	UNP B0S4P3
C	0	HIS	-	EXPRESSION TAG	UNP B0S4P3

- Molecule 2 is water.

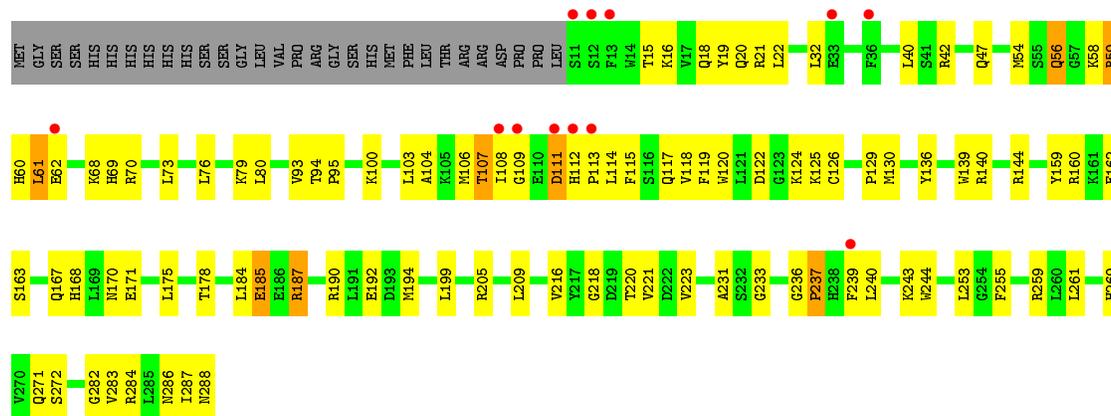
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	125	Total O 125 125	0	0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
2	B	123	Total 123	O 123	0	0
2	C	168	Total 168	O 168	0	0





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	124.92Å 128.56Å 124.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.67 – 2.50 44.67 – 2.09	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.67-2.50) 99.5 (44.67-2.09)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.93 (at 2.08Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.214 , 0.275 0.214 , 0.275	Depositor DCC
$R_{free}$ test set	1768 reflections (5.07%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.9	Xtriage
Anisotropy	0.571	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 52.2	EDS
Estimated twinning fraction	0.026 for -k,-h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	2 of 59243 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7189	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.80% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.37	0/2303	0.61	0/3105
1	B	0.36	0/2311	0.62	1/3116 (0.0%)
1	C	0.38	0/2303	0.63	0/3105
All	All	0.37	0/6917	0.62	1/9326 (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	109	GLY	N-CA-C	-5.31	99.83	113.10

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	2255	0	2249	97	0
1	B	2263	0	2260	126	0
1	C	2255	0	2249	105	0
2	A	125	0	0	13	0
2	B	123	0	0	7	0
2	C	168	0	0	13	0
All	All	7189	0	6758	324	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 24.

The worst 5 of 324 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:C:104:ALA:HA	1:C:108:ILE:HB	1.32	1.10
1:B:108:ILE:HG12	1:B:109:GLY:H	1.00	1.06
1:C:107:THR:HG23	1:C:240:LEU:HD21	1.33	1.06
1:C:108:ILE:HG12	1:C:112:HIS:HB3	1.44	1.00
1:C:108:ILE:HD11	1:C:114:LEU:HB3	1.46	0.97

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	276/308 (90%)	249 (90%)	21 (8%)	6 (2%)	8	13
1	B	277/308 (90%)	253 (91%)	19 (7%)	5 (2%)	11	18
1	C	276/308 (90%)	264 (96%)	9 (3%)	3 (1%)	17	31
All	All	829/924 (90%)	766 (92%)	49 (6%)	14 (2%)	11	19

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	266	GLY
1	B	108	ILE
1	B	107	THR
1	B	218	GLY
1	B	286	ASN

### 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	242/269 (90%)	225 (93%)	17 (7%)	19	34
1	B	243/269 (90%)	222 (91%)	21 (9%)	13	24
1	C	242/269 (90%)	230 (95%)	12 (5%)	30	53
All	All	727/807 (90%)	677 (93%)	50 (7%)	19	35

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	67	VAL
1	B	118	VAL
1	C	144	ARG
1	B	70	ARG
1	B	103	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	47	GLN
1	B	271	GLN
1	C	167	GLN
1	B	167	GLN
1	B	238	HIS

### 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 [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	278/308 (90%)	0.27	32 (11%) 6 6	17, 37, 77, 86	0
1	B	279/308 (90%)	0.38	34 (12%) 5 5	14, 36, 78, 88	0
1	C	278/308 (90%)	-0.13	12 (4%) 39 44	11, 34, 66, 78	0
All	All	835/924 (90%)	0.17	78 (9%) 11 11	11, 36, 76, 88	0

The worst 5 of 78 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	216	VAL	18.8
1	B	215	VAL	17.9
1	B	218	GLY	11.7
1	B	165	GLY	11.5
1	C	109	GLY	9.6

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.