



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 11:02 PM GMT

PDB ID : 1W9H  
Title : THE STRUCTURE OF A PIWI PROTEIN FROM ARCHAEOGLOBUS FULGIDUS.  
Authors : Parker, J.S.; Roe, S.M.; Barford, D.  
Deposited on : 2004-10-13  
Resolution : 1.95 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  
<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)  
Xtriage (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

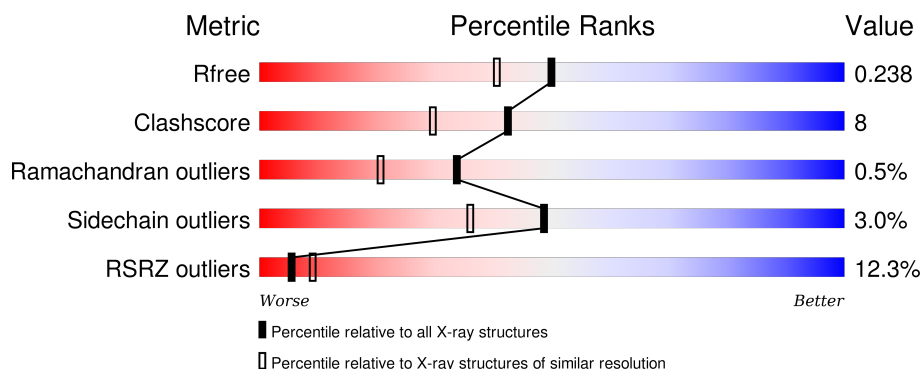
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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	427	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 3482 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 HYPOTHETICAL PROTEIN AF1318.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	398	Total	C	N	O	S	0	0	0
			3233	2097	531	596	9			

- Molecule 2 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Cd	0	0
			2	2		

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	5	Total	Ni	0	0
			5	5		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	4	Total	Cl	0	0
			4	4		

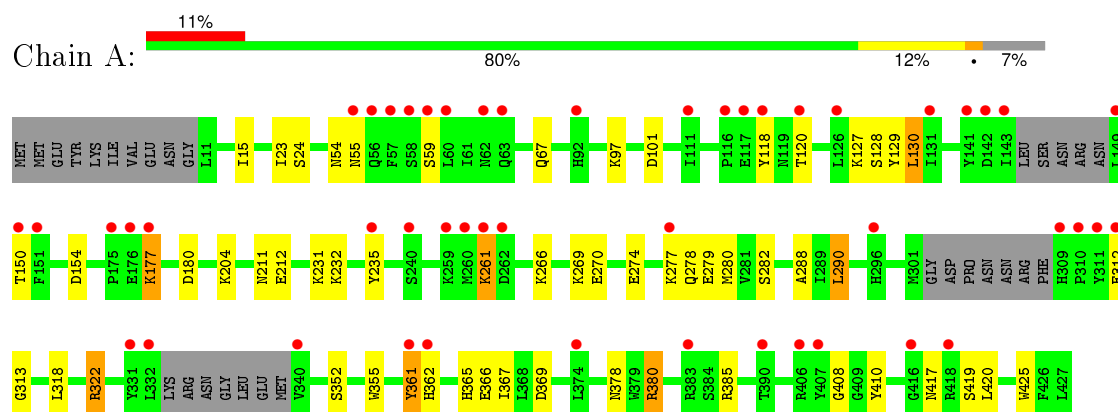
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	238	Total	O	0	0
			238	238		

### 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 errors 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: HYPOTHETICAL PROTEIN AF1318



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.09 Å 137.86 Å 51.66 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	69.01 – 1.95 30.88 – 1.95	Depositor EDS
% Data completeness (in resolution range)	98.5 (69.01-1.95) 98.5 (30.88-1.95)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.48 (at 1.95 Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.192 , 0.237 0.191 , 0.238	Depositor DCC
$R_{free}$ test set	1813 reflections (5.01%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.6	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.44 , 55.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 36197 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3482	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.15% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NI, CL, CD

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	0.83	2/3314 (0.1%)	0.78	1/4498 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	55	ASN	C-N	-6.91	1.18	1.34
1	A	180	ASP	CB-CG	5.20	1.62	1.51

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	130	LEU	CA-CB-CG	-7.39	98.29	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	3233	0	3242	54	0
2	A	2	0	0	0	0
3	A	5	0	0	0	0
4	A	4	0	0	0	0
5	A	238	0	0	11	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	3482	0	3242	54	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 8.

All (54) 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:362:HIS:CE1	1:A:366:GLU:OE2	2.03	1.11
1:A:290:LEU:HD21	1:A:367:ILE:HG21	1.40	1.00
1:A:261:LYS:HA	5:A:2156:HOH:O	1.67	0.94
1:A:378:ASN:ND2	5:A:2208:HOH:O	2.03	0.90
1:A:235:TYR:HD1	1:A:280:MET:CE	1.93	0.81
1:A:290:LEU:HD21	1:A:367:ILE:CG2	2.16	0.74
1:A:235:TYR:HD1	1:A:280:MET:HE1	1.53	0.74
1:A:211:ASN:HD21	1:A:425:TRP:HE1	1.34	0.73
1:A:154:ASP:OD1	5:A:2089:HOH:O	2.06	0.71
1:A:385:ARG:N	5:A:2216:HOH:O	2.22	0.70
1:A:355:TRP:HH2	1:A:367:ILE:HD12	1.61	0.66
1:A:266:LYS:O	1:A:270:GLU:HG3	1.96	0.66
1:A:322:ARG:HD2	5:A:2181:HOH:O	1.95	0.66
1:A:318:LEU:HB2	1:A:322:ARG:HG2	1.77	0.65
1:A:362:HIS:NE2	1:A:366:GLU:OE2	2.31	0.61
1:A:367:ILE:HG23	5:A:2150:HOH:O	2.01	0.61
1:A:312:GLU:CG	1:A:313:GLY:N	2.61	0.60
1:A:378:ASN:ND2	1:A:385:ARG:HB2	2.18	0.59
1:A:235:TYR:HD1	1:A:280:MET:HE2	1.67	0.59
1:A:355:TRP:CZ2	1:A:367:ILE:HD11	2.38	0.59
1:A:235:TYR:CD1	1:A:280:MET:HE1	2.35	0.59
1:A:312:GLU:HG2	1:A:313:GLY:N	2.18	0.58
1:A:235:TYR:HE1	1:A:278:GLN:NE2	2.01	0.58
1:A:355:TRP:CH2	1:A:367:ILE:HD12	2.39	0.56
1:A:235:TYR:CD1	1:A:280:MET:CE	2.83	0.56
1:A:177:LYS:HG3	5:A:2111:HOH:O	2.05	0.55
1:A:235:TYR:CE1	1:A:278:GLN:NE2	2.76	0.54
1:A:67:GLN:NE2	1:A:150:THR:HG22	2.24	0.53
1:A:127:LYS:HE2	5:A:2075:HOH:O	2.09	0.53
1:A:355:TRP:HZ2	1:A:367:ILE:HD11	1.73	0.53
1:A:355:TRP:CH2	1:A:367:ILE:CD1	2.92	0.52
1:A:365:HIS:CE1	1:A:369:ASP:OD1	2.63	0.52
1:A:277:LYS:C	1:A:279:GLU:H	2.14	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:417:ASN:OD1	1:A:419:SER:HB3	2.12	0.50
1:A:312:GLU:HG2	1:A:313:GLY:H	1.76	0.49
1:A:235:TYR:CD1	1:A:280:MET:HE2	2.47	0.48
1:A:231:LYS:NZ	1:A:274:GLU:OE1	2.32	0.48
1:A:212:GLU:HG3	1:A:232:LYS:HG2	1.96	0.48
1:A:269:LYS:HE2	1:A:352:SER:OG	2.15	0.47
1:A:355:TRP:HH2	1:A:367:ILE:CD1	2.25	0.46
1:A:378:ASN:OD1	1:A:380:ARG:HB2	2.16	0.45
1:A:211:ASN:ND2	1:A:425:TRP:HE1	2.09	0.45
1:A:378:ASN:HD22	1:A:385:ARG:HB2	1.82	0.45
1:A:235:TYR:HA	1:A:280:MET:CE	2.47	0.44
1:A:361:TYR:HB3	5:A:2200:HOH:O	2.18	0.43
1:A:128:SER:HB2	1:A:410:TYR:O	2.18	0.43
1:A:23:ILE:O	1:A:24:SER:HB2	2.20	0.42
1:A:97:LYS:HD2	1:A:129:TYR:CZ	2.55	0.41
1:A:385:ARG:NH2	5:A:2215:HOH:O	2.27	0.41
1:A:288:ALA:HB2	1:A:355:TRP:CZ2	2.55	0.41
1:A:177:LYS:HB3	1:A:204:LYS:HD2	2.03	0.41
1:A:408:GLY:HA2	5:A:2224:HOH:O	2.20	0.40
1:A:97:LYS:NZ	1:A:101:ASP:OD2	2.49	0.40
1:A:420:LEU:HD23	1:A:420:LEU:HA	1.92	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	390/427 (91%)	371 (95%)	17 (4%)	2 (0%)	34 21

All (2) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	A	261	LYS
1	A	15	ILE

### 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	366/394 (93%)	355 (97%)	11 (3%)	48 36

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	54	ASN
1	A	59	SER
1	A	118	TYR
1	A	120	THR
1	A	130	LEU
1	A	177	LYS
1	A	282	SER
1	A	290	LEU
1	A	322	ARG
1	A	361	TYR
1	A	380	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	54	ASN
1	A	211	ASN
1	A	278	GLN
1	A	362	HIS
1	A	365	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](#)

Of 11 ligands modelled in this entry, 11 are 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 [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	398/427 (93%)	0.78	49 (12%) 5 9	-3, 13, 43, 69	0

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	118	TYR	10.5
1	A	151	PHE	8.1
1	A	57	PHE	6.9
1	A	361	TYR	6.7
1	A	143	ILE	6.2
1	A	261	LYS	6.0
1	A	59	SER	5.4
1	A	332	LEU	4.9
1	A	416	GLY	4.7
1	A	331	TYR	4.6
1	A	260	MET	4.4
1	A	340	VAL	4.2
1	A	63	GLN	4.0
1	A	262	ASP	4.0
1	A	60	LEU	4.0
1	A	259	LYS	3.9
1	A	117	GLU	3.8
1	A	116	PRO	3.7
1	A	150	THR	3.4
1	A	149	LEU	3.3
1	A	311	TYR	3.2
1	A	120	THR	3.1
1	A	310	PRO	3.0
1	A	407	TYR	3.0
1	A	58	SER	3.0
1	A	309	HIS	2.9
1	A	142	ASP	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	418	ARG	2.8
1	A	383	ARG	2.7
1	A	296	HIS	2.6
1	A	240	SER	2.6
1	A	131	ILE	2.6
1	A	126	LEU	2.5
1	A	374	LEU	2.5
1	A	62	ASN	2.5
1	A	175	PRO	2.5
1	A	55	ASN	2.5
1	A	277	LYS	2.5
1	A	141	TYR	2.4
1	A	406	ARG	2.4
1	A	92	HIS	2.3
1	A	56	GLN	2.3
1	A	176	GLU	2.2
1	A	312	GLU	2.2
1	A	177	LYS	2.1
1	A	390	THR	2.1
1	A	111	ILE	2.1
1	A	235	TYR	2.1
1	A	362	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](#)

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
4	CL	A	1436	1/1	0.99	0.11	-1.55	13,13,13,13	0
2	CD	A	1428	1/1	1.00	0.08	-3.19	8,8,8,8	0
2	CD	A	1429	1/1	0.99	0.05	-4.06	14,14,14,14	0
3	NI	A	1432	1/1	0.98	0.04	-	26,26,26,26	0
4	CL	A	1437	1/1	0.97	0.07	-	15,15,15,15	0
4	CL	A	1435	1/1	0.99	0.05	-	15,15,15,15	0
4	CL	A	1438	1/1	0.97	0.09	-	17,17,17,17	0
3	NI	A	1433	1/1	0.95	0.06	-	35,35,35,35	0
3	NI	A	1430	1/1	0.98	0.06	-	13,13,13,13	0
3	NI	A	1434	1/1	0.99	0.08	-	27,27,27,27	0
3	NI	A	1431	1/1	0.98	0.10	-	39,39,39,39	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.