



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 03:47 pm BST

PDB ID : 4FZ2
Title : Crystal structure of the fourth type of archaeal tRNA splicing endonuclease from *Candidatus Micrarchaeum acidiphilum* ARMAN-2
Authors : Hirata, A.; Fujishima, K.; Yamagami, R.; Kawamura, T.; Banfield, J.F.; Kanai, A.; Hori, H.
Deposited on : 2012-07-06
Resolution : 2.25 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

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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

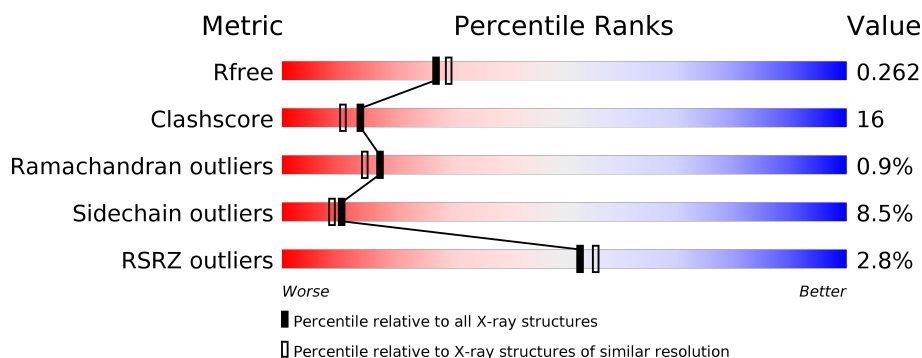
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.25 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}	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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 ≥ 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 $\leq 5\%$. 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	395	<div> <div>2%</div> <div> <div></div> <div>68%</div> <div>25%</div> <div>5%</div> </div> </div>
1	B	395	<div> <div>3%</div> <div> <div></div> <div>68%</div> <div>25%</div> <div>5%</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6456 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 tRNA intron endonuclease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	386	Total	C	N	O	S	0	0	0
			3152	2027	545	572	8			
1	B	386	Total	C	N	O	S	0	0	0
			3152	2027	545	572	8			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	388	LEU	-	EXPRESSION TAG	UNP C7DIA5
A	389	GLU	-	EXPRESSION TAG	UNP C7DIA5
A	390	HIS	-	EXPRESSION TAG	UNP C7DIA5
A	391	HIS	-	EXPRESSION TAG	UNP C7DIA5
A	392	HIS	-	EXPRESSION TAG	UNP C7DIA5
A	393	HIS	-	EXPRESSION TAG	UNP C7DIA5
A	394	HIS	-	EXPRESSION TAG	UNP C7DIA5
A	395	HIS	-	EXPRESSION TAG	UNP C7DIA5
B	388	LEU	-	EXPRESSION TAG	UNP C7DIA5
B	389	GLU	-	EXPRESSION TAG	UNP C7DIA5
B	390	HIS	-	EXPRESSION TAG	UNP C7DIA5
B	391	HIS	-	EXPRESSION TAG	UNP C7DIA5
B	392	HIS	-	EXPRESSION TAG	UNP C7DIA5
B	393	HIS	-	EXPRESSION TAG	UNP C7DIA5
B	394	HIS	-	EXPRESSION TAG	UNP C7DIA5
B	395	HIS	-	EXPRESSION TAG	UNP C7DIA5

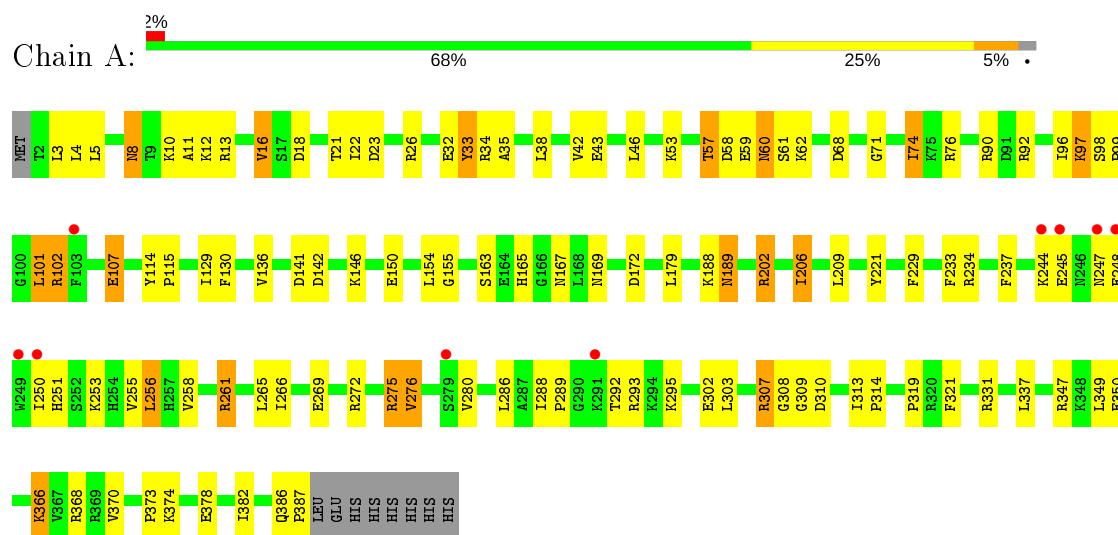
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	81	Total	O	0	0
			81	81		
2	B	71	Total	O	0	0
			71	71		

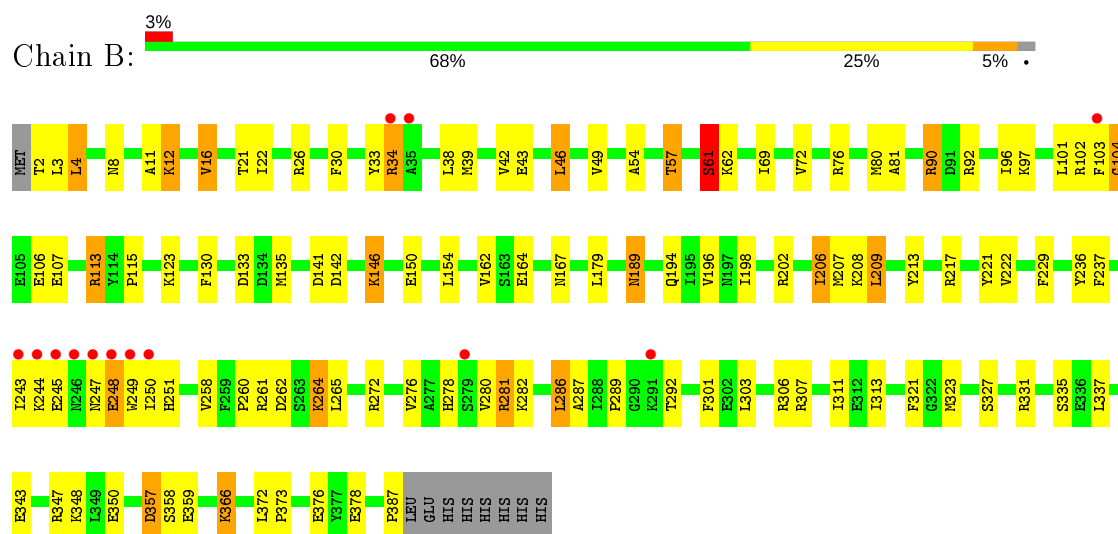
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: tRNA intron endonuclease



• Molecule 1: tRNA intron endonuclease



4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	112.02Å 112.02Å 81.08Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.41 – 2.25 37.40 – 2.25	Depositor EDS
% Data completeness (in resolution range)	92.4 (37.41-2.25) 99.4 (37.40-2.25)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.22 (at 2.24Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.217 , 0.257 0.221 , 0.262	Depositor DCC
R_{free} test set	2722 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	41.3	Xtriage
Anisotropy	0.443	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 31.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.017 for -h,-k,l 0.488 for h,-h-k,-l 0.019 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6456	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.46% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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	0.33	0/3221	0.52	0/4338
1	B	0.34	0/3221	0.53	0/4338
All	All	0.34	0/6442	0.52	0/8676

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	3152	0	3191	102	0
1	B	3152	0	3191	106	0
2	A	81	0	0	5	0
2	B	71	0	0	4	0
All	All	6456	0	6382	200	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 200 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:10:LYS:HA	1:A:11:ALA:HB2	1.39	1.04

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:4:LEU:H	1:A:21:THR:HG21	1.33	0.94
1:B:57:THR:HG23	1:B:61:SER:HA	1.51	0.92
1:A:189:ASN:HD22	1:A:189:ASN:H	1.19	0.89
1:B:208:LYS:HE3	1:B:260:PRO:HB3	1.55	0.89

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	384/395 (97%)	361 (94%)	21 (6%)	2 (0%)	29	29
1	B	384/395 (97%)	364 (95%)	15 (4%)	5 (1%)	12	8
All	All	768/790 (97%)	725 (94%)	36 (5%)	7 (1%)	17	14

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	357	ASP
1	B	164	GLU
1	A	308	GLY
1	B	104	GLY
1	B	248	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	340/349 (97%)	309 (91%)	31 (9%)	9	7
1	B	340/349 (97%)	313 (92%)	27 (8%)	12	10
All	All	680/698 (97%)	622 (92%)	58 (8%)	10	9

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	307	ARG
1	B	12	LYS
1	B	281	ARG
1	A	337	LEU
1	A	366	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	189	ASN
1	B	247	ASN
1	B	110	GLN
1	A	110	GLN
1	A	246	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 ⓘ

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, 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	386/395 (97%)	0.01	9 (2%) 60 63	29, 45, 80, 122	0
1	B	386/395 (97%)	0.05	13 (3%) 45 47	29, 44, 79, 126	0
All	All	772/790 (97%)	0.03	22 (2%) 53 55	29, 45, 80, 126	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	244	LYS	11.0
1	B	244	LYS	10.0
1	A	249	TRP	9.3
1	B	249	TRP	7.9
1	B	250	ILE	5.5

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.