



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 7, 2024 – 11:01 PM EDT

PDB ID : 3LRX  
Title : Crystal Structure of the C-terminal domain (residues 78-226) of PF1911 hydrogenase from *Pyrococcus furiosus*, Northeast Structural Genomics Consortium Target PfR246A  
Authors : Forouhar, F.; Abashidze, M.; Seetharaman, J.; Mao, M.; Xiao, R.; Ciccocanti, C.; Foote, E.L.; Belote, R.L.; Everett, J.K.; Nair, R.; Acton, T.B.; Rost, B.; Montelione, G.T.; Tong, L.; Hunt, J.F.; Northeast Structural Genomics Consortium (NESG)  
Deposited on : 2010-02-11  
Resolution : 2.60 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)

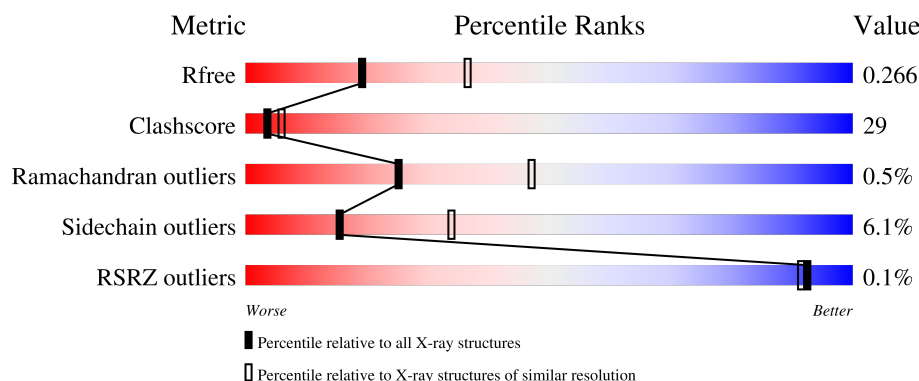
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	164625	3775 (2.60-2.60)
Clashscore	180529	4181 (2.60-2.60)
Ramachandran outliers	177936	4129 (2.60-2.60)
Sidechain outliers	177891	4129 (2.60-2.60)
RSRZ outliers	164620	3775 (2.60-2.60)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	158	
1	B	158	
1	C	158	

*Continued on next page...*

Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	D	158	<div><div></div><div>43%37%17%</div></div>
1	E	158	<div><div>%</div><div></div><div>44%33%18%</div></div>
1	F	158	<div><div></div><div>44%35%20%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6291 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 hydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	132	Total	C	N	O	Se	0	0	0
			1045	679	168	193	5			
1	B	132	Total	C	N	O	Se	0	0	0
			1045	679	168	193	5			
1	C	129	Total	C	N	O	Se	0	0	0
			1026	669	165	187	5			
1	D	131	Total	C	N	O	Se	0	0	0
			1037	673	167	192	5			
1	E	129	Total	C	N	O	Se	0	0	0
			1026	671	167	183	5			
1	F	127	Total	C	N	O	Se	0	0	0
			1005	655	160	185	5			

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	77	MSE	-	initiating methionine	UNP Q8TZS3
A	227	LEU	-	expression tag	UNP Q8TZS3
A	228	GLU	-	expression tag	UNP Q8TZS3
A	229	HIS	-	expression tag	UNP Q8TZS3
A	230	HIS	-	expression tag	UNP Q8TZS3
A	231	HIS	-	expression tag	UNP Q8TZS3
A	232	HIS	-	expression tag	UNP Q8TZS3
A	233	HIS	-	expression tag	UNP Q8TZS3
A	234	HIS	-	expression tag	UNP Q8TZS3
B	77	MSE	-	initiating methionine	UNP Q8TZS3
B	227	LEU	-	expression tag	UNP Q8TZS3
B	228	GLU	-	expression tag	UNP Q8TZS3
B	229	HIS	-	expression tag	UNP Q8TZS3
B	230	HIS	-	expression tag	UNP Q8TZS3
B	231	HIS	-	expression tag	UNP Q8TZS3
B	232	HIS	-	expression tag	UNP Q8TZS3
B	233	HIS	-	expression tag	UNP Q8TZS3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	234	HIS	-	expression tag	UNP Q8TZS3
C	77	MSE	-	initiating methionine	UNP Q8TZS3
C	227	LEU	-	expression tag	UNP Q8TZS3
C	228	GLU	-	expression tag	UNP Q8TZS3
C	229	HIS	-	expression tag	UNP Q8TZS3
C	230	HIS	-	expression tag	UNP Q8TZS3
C	231	HIS	-	expression tag	UNP Q8TZS3
C	232	HIS	-	expression tag	UNP Q8TZS3
C	233	HIS	-	expression tag	UNP Q8TZS3
C	234	HIS	-	expression tag	UNP Q8TZS3
D	77	MSE	-	initiating methionine	UNP Q8TZS3
D	227	LEU	-	expression tag	UNP Q8TZS3
D	228	GLU	-	expression tag	UNP Q8TZS3
D	229	HIS	-	expression tag	UNP Q8TZS3
D	230	HIS	-	expression tag	UNP Q8TZS3
D	231	HIS	-	expression tag	UNP Q8TZS3
D	232	HIS	-	expression tag	UNP Q8TZS3
D	233	HIS	-	expression tag	UNP Q8TZS3
D	234	HIS	-	expression tag	UNP Q8TZS3
E	77	MSE	-	initiating methionine	UNP Q8TZS3
E	227	LEU	-	expression tag	UNP Q8TZS3
E	228	GLU	-	expression tag	UNP Q8TZS3
E	229	HIS	-	expression tag	UNP Q8TZS3
E	230	HIS	-	expression tag	UNP Q8TZS3
E	231	HIS	-	expression tag	UNP Q8TZS3
E	232	HIS	-	expression tag	UNP Q8TZS3
E	233	HIS	-	expression tag	UNP Q8TZS3
E	234	HIS	-	expression tag	UNP Q8TZS3
F	77	MSE	-	initiating methionine	UNP Q8TZS3
F	227	LEU	-	expression tag	UNP Q8TZS3
F	228	GLU	-	expression tag	UNP Q8TZS3
F	229	HIS	-	expression tag	UNP Q8TZS3
F	230	HIS	-	expression tag	UNP Q8TZS3
F	231	HIS	-	expression tag	UNP Q8TZS3
F	232	HIS	-	expression tag	UNP Q8TZS3
F	233	HIS	-	expression tag	UNP Q8TZS3
F	234	HIS	-	expression tag	UNP Q8TZS3

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	21	Total O 21 21	0	0

*Continued on next page...*

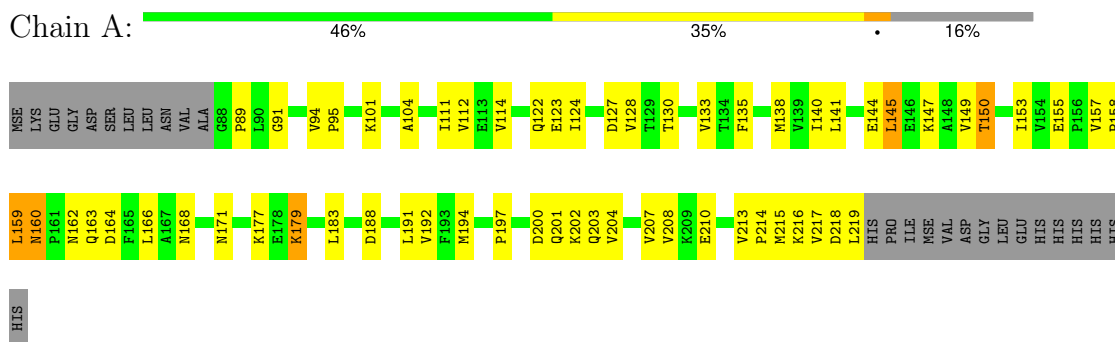
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	28	Total 28	O 28	0	0
2	C	17	Total 17	O 17	0	0
2	D	21	Total 21	O 21	0	0
2	E	12	Total 12	O 12	0	0
2	F	8	Total 8	O 8	0	0

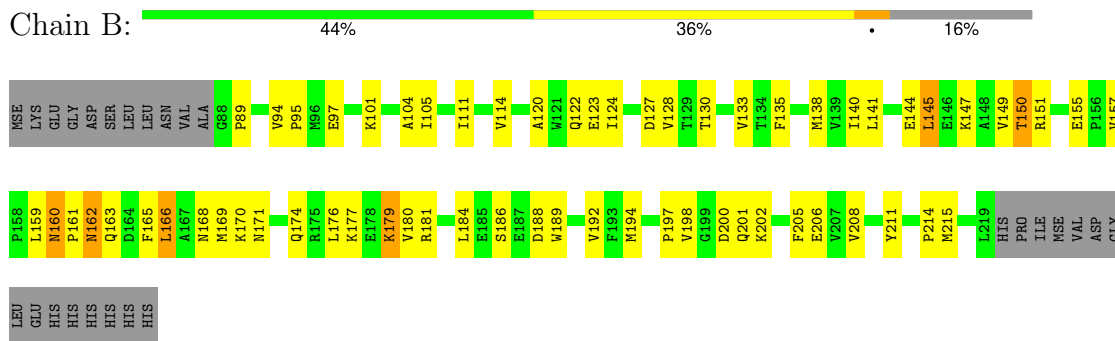
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Putative hydrogenase



#### • Molecule 1: Putative hydrogenase



#### • Molecule 1: Putative hydrogenase



#### • Molecule 1: Putative hydrogenase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.20Å 77.20Å 110.04Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.30 – 2.60 19.30 – 2.60	Depositor EDS
% Data completeness (in resolution range)	78.6 (19.30-2.60) 92.9 (19.30-2.60)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.33 (at 2.61Å)	Xtriage
Refinement program	CNS 1.2 & XtalView, REFMAC	Depositor
R, $R_{free}$	0.195 , 0.252 0.214 , 0.266	Depositor DCC
$R_{free}$ test set	2091 reflections (9.35%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.5	Xtriage
Anisotropy	0.212	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 59.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.025 for -h,-k,l 0.047 for h,-h-k,-l 0.036 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6291	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.61 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.5072e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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](#)

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	0.47	0/1063	0.62	0/1436
1	B	0.45	0/1063	1.07	3/1436 (0.2%)
1	C	0.43	0/1043	0.63	0/1407
1	D	0.45	0/1055	0.61	0/1425
1	E	0.43	0/1045	0.63	1/1411 (0.1%)
1	F	0.43	0/1021	0.63	0/1376
All	All	0.45	0/6290	0.72	4/8491 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	151	ARG	NE-CZ-NH1	-22.92	108.84	120.30
1	B	151	ARG	NE-CZ-NH2	22.49	131.54	120.30
1	B	151	ARG	CD-NE-CZ	11.44	139.61	123.60
1	E	151	ARG	NE-CZ-NH2	-5.12	117.74	120.30

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	1045	0	1069	63	0
1	B	1045	0	1069	70	0
1	C	1026	0	1055	52	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1037	0	1058	70	0
1	E	1026	0	1059	65	0
1	F	1005	0	1030	63	0
2	A	21	0	0	1	0
2	B	28	0	0	2	0
2	C	17	0	0	1	0
2	D	21	0	0	5	0
2	E	12	0	0	2	0
2	F	8	0	0	1	0
All	All	6291	0	6340	359	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 29.

The worst 5 of 359 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:F:192:VAL:HG11	1:F:208:VAL:HG21	1.45	0.96
1:B:192:VAL:HG11	1:B:208:VAL:HG21	1.49	0.95
1:D:160:ASN:HD22	1:D:163:GLN:HG3	1.33	0.94
1:A:194:MSE:HE2	1:A:201:GLN:HG2	1.54	0.90
1:A:127:ASP:OD1	1:A:150:THR:HG21	1.72	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	130/158 (82%)	117 (90%)	13 (10%)	0	100	100
1	B	130/158 (82%)	117 (90%)	11 (8%)	2 (2%)	8	18

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	125/158 (79%)	115 (92%)	10 (8%)	0	100	100
1	D	129/158 (82%)	115 (89%)	13 (10%)	1 (1%)	16	34
1	E	125/158 (79%)	112 (90%)	13 (10%)	0	100	100
1	F	123/158 (78%)	112 (91%)	10 (8%)	1 (1%)	16	34
All	All	762/948 (80%)	688 (90%)	70 (9%)	4 (0%)	25	47

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	186	SER
1	B	174	GLN
1	F	218	ASP
1	D	161	PRO

### 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	117/133 (88%)	109 (93%)	8 (7%)	13	28
1	B	117/133 (88%)	108 (92%)	9 (8%)	10	22
1	C	115/133 (86%)	108 (94%)	7 (6%)	15	34
1	D	116/133 (87%)	109 (94%)	7 (6%)	16	35
1	E	115/133 (86%)	107 (93%)	8 (7%)	12	27
1	F	112/133 (84%)	109 (97%)	3 (3%)	40	66
All	All	692/798 (87%)	650 (94%)	42 (6%)	15	34

5 of 42 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	176	LEU
1	E	179	LYS
1	D	183	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	138	MSE
1	E	195	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	122	GLN
1	D	201	GLN
1	F	171	ASN
1	D	163	GLN
1	D	203	GLN

### 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 oligosaccharides 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(Å²)	Q<0.9	
1	A	127/158 (80%)	-0.32	0	100	100	12, 30, 51, 60	0
1	B	127/158 (80%)	-0.26	0	100	100	12, 33, 52, 65	0
1	C	124/158 (78%)	-0.16	0	100	100	19, 36, 56, 65	0
1	D	126/158 (79%)	-0.12	0	100	100	13, 35, 61, 79	0
1	E	124/158 (78%)	-0.00	1 (0%)	82	79	18, 39, 59, 70	0
1	F	122/158 (77%)	-0.25	0	100	100	16, 33, 54, 67	0
All	All	750/948 (79%)	-0.19	1 (0%)	92	91	12, 34, 56, 79	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	184	LEU	2.3

### 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 monosaccharides 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.