



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 18, 2024 – 08:57 PM EST

PDB ID : 4H6Y  
Title : Crystal structure of the DH-PH-PH domain of FARP1  
Authors : He, X.; Zhang, X.  
Deposited on : 2012-09-19  
Resolution : 4.09 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  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

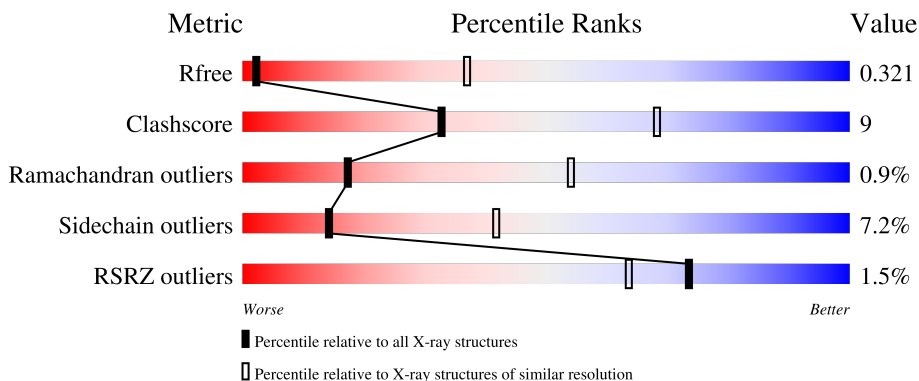
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.09 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	1193 (4.50-3.70)
Clashscore	141614	1003 (4.44-3.76)
Ramachandran outliers	138981	1005 (4.48-3.72)
Sidechain outliers	138945	1199 (4.50-3.70)
RSRZ outliers	127900	1034 (4.50-3.70)

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	501	
1	B	501	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 5955 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 FERM, RhoGEF and pleckstrin domain-containing protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	402	3005	1949	499	537	20	0	0	0
1	B	408	2950	1903	496	532	19	0	0	0

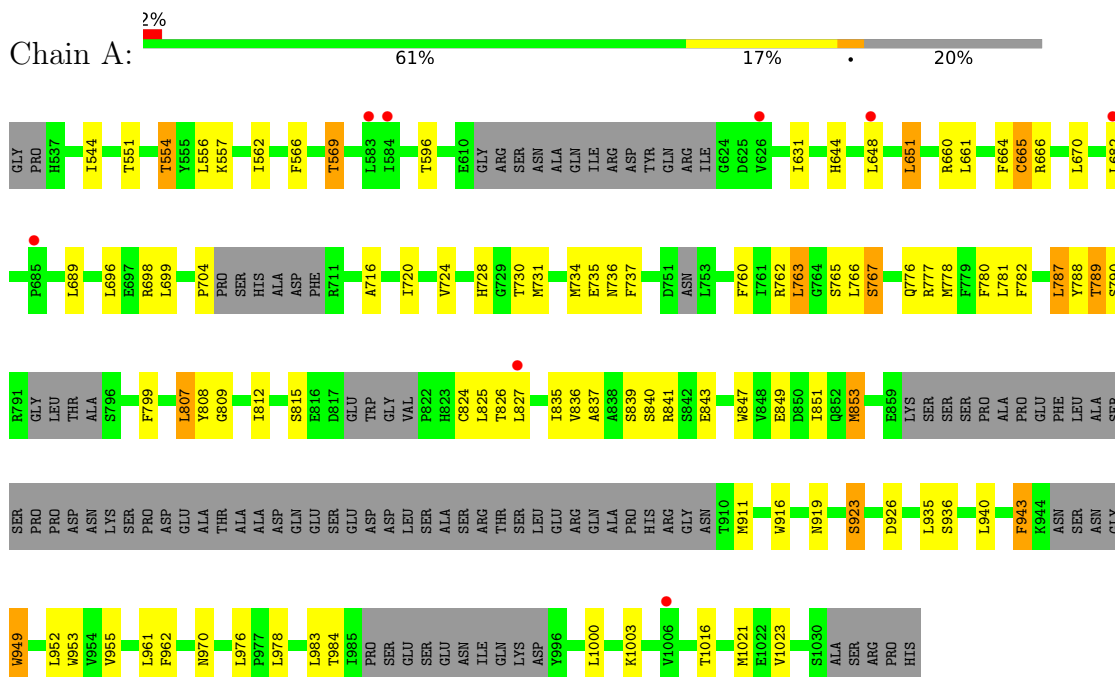
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	535	GLY	-	expression tag	UNP Q9Y4F1
A	536	PRO	-	expression tag	UNP Q9Y4F1
A	537	HIS	-	expression tag	UNP Q9Y4F1
A	538	MET	-	expression tag	UNP Q9Y4F1
B	535	GLY	-	expression tag	UNP Q9Y4F1
B	536	PRO	-	expression tag	UNP Q9Y4F1
B	537	HIS	-	expression tag	UNP Q9Y4F1
B	538	MET	-	expression tag	UNP Q9Y4F1

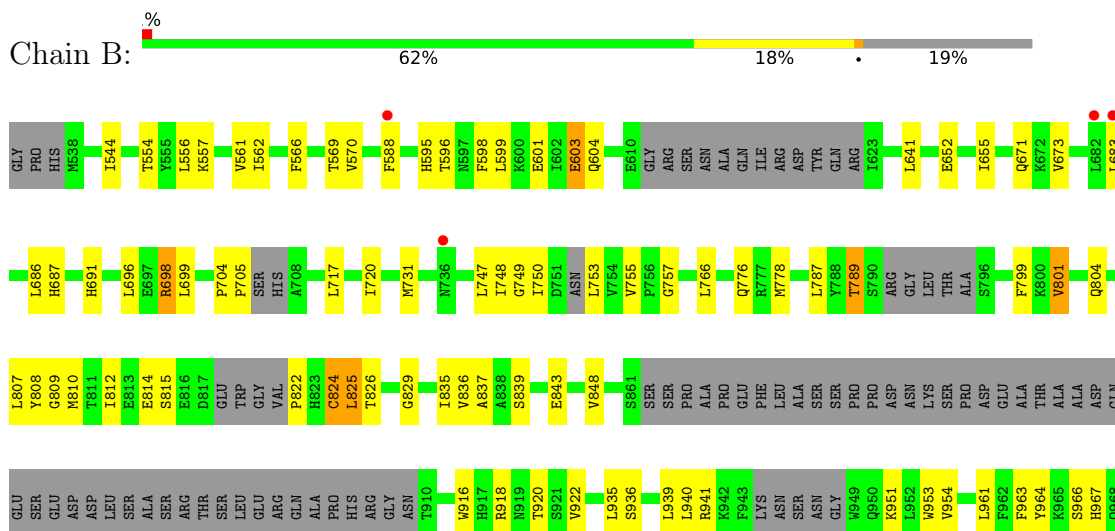
### 3 Residue-property plots [i](#)

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: FERM, RhoGEF and pleckstrin domain-containing protein 1



- Molecule 1: FERM, RhoGEF and pleckstrin domain-containing protein 1



D969	D970	L976	L979	T985	P986	SER	GLU	SER	GLU	ASN	ILE	GLN	K994	F998	F1009	Y1015	T1016	F1017	E1018	R1019	W1020	M1021	T1028	S1029	SER	ALA	SER	ARG	PRO	HIS
------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	142.24Å 142.24Å 105.18Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.57 – 4.09 42.57 – 4.09	Depositor EDS
% Data completeness (in resolution range)	96.0 (42.57-4.09) 95.8 (42.57-4.09)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.60 (at 4.13Å)	Xtrriage
Refinement program	PHENIX 1.7.2_869	Depositor
R, $R_{free}$	0.289 , 0.317 0.284 , 0.321	Depositor DCC
$R_{free}$ test set	949 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	145.9	Xtrriage
Anisotropy	0.075	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 155.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.029 for -h,-k,l 0.448 for h,-h-k,-l 0.033 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	5955	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	176.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.60% 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 [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	0.26	0/3070	0.44	1/4178 (0.0%)
1	B	0.26	0/3013	0.43	2/4112 (0.0%)
All	All	0.26	0/6083	0.44	3/8290 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	822	PRO	N-CA-CB	6.13	110.66	103.30
1	B	986	PRO	N-CA-CB	5.83	110.29	103.30
1	A	704	PRO	N-CA-CB	5.62	110.04	103.30

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	3005	0	2704	51	0
1	B	2950	0	2564	47	0
All	All	5955	0	5268	98	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 9.

The worst 5 of 98 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:B:961:LEU:HB2	1:B:976:LEU:HB3	1.70	0.74
1:A:760:PHE:HZ	1:A:763:LEU:HD12	1.53	0.73
1:B:810:MET:HB2	1:B:829:GLY:HA2	1.73	0.70
1:A:631:ILE:HD12	1:A:720:ILE:HD11	1.72	0.70
1:A:765:SER:HB3	1:A:919:ASN:HA	1.74	0.69

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	384/501 (77%)	349 (91%)	32 (8%)	3 (1%)	19	58
1	B	390/501 (78%)	357 (92%)	29 (7%)	4 (1%)	15	52
All	All	774/1002 (77%)	706 (91%)	61 (8%)	7 (1%)	17	54

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	767	SER
1	B	807	LEU
1	A	807	LEU
1	A	1003	LYS
1	B	1028	THR

#### 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	282/449 (63%)	258 (92%)	24 (8%)	10	37
1	B	263/449 (59%)	248 (94%)	15 (6%)	20	49
All	All	545/898 (61%)	506 (93%)	39 (7%)	14	42

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	698	ARG
1	B	941	ARG
1	B	789	THR
1	B	825	LEU
1	B	970	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 monosaccharides 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

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	402/501 (80%)	-0.02	8 (1%) 65 56	114, 173, 226, 264	0
1	B	408/501 (81%)	-0.07	4 (0%) 82 74	108, 172, 231, 277	0
All	All	810/1002 (80%)	-0.05	12 (1%) 73 63	108, 172, 229, 277	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	588	PHE	2.3
1	A	685	PRO	2.3
1	A	626	VAL	2.2
1	A	827	LEU	2.2
1	A	648	LEU	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 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.