



# Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 12:07 pm BST

PDB ID : 5CVL  
Title : WDR48 (UAF-1), residues 2-580  
Authors : HARRIS, S.F.; YIN, J.  
Deposited on : 2015-07-27  
Resolution : 3.00 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

<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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
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

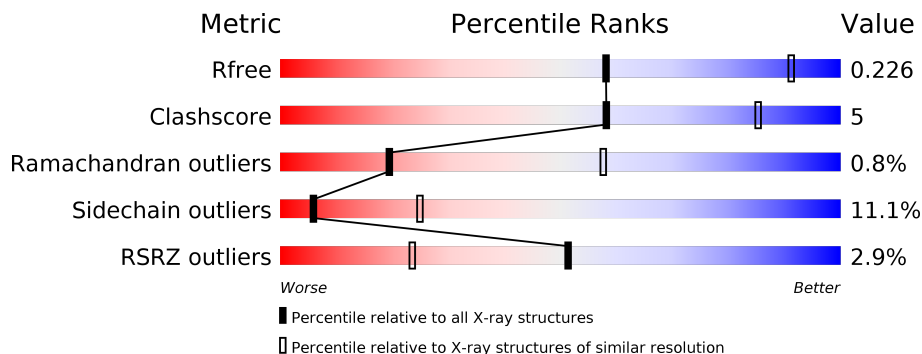
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	598	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	PO4	A	602	-	-	-	X

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4292 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 WD repeat-containing protein 48.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	526	4138	2611	730	774	23	0	0	0

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-14	MET	-	initiating methionine	UNP Q8TAF3
A	-13	HIS	-	expression tag	UNP Q8TAF3
A	-12	HIS	-	expression tag	UNP Q8TAF3
A	-11	HIS	-	expression tag	UNP Q8TAF3
A	-10	HIS	-	expression tag	UNP Q8TAF3
A	-9	HIS	-	expression tag	UNP Q8TAF3
A	-8	HIS	-	expression tag	UNP Q8TAF3
A	-7	GLY	-	expression tag	UNP Q8TAF3
A	-6	GLU	-	expression tag	UNP Q8TAF3
A	-5	ASN	-	expression tag	UNP Q8TAF3
A	-4	LEU	-	expression tag	UNP Q8TAF3
A	-3	TYR	-	expression tag	UNP Q8TAF3
A	-2	PHE	-	expression tag	UNP Q8TAF3
A	-1	GLN	-	expression tag	UNP Q8TAF3
A	0	GLY	-	expression tag	UNP Q8TAF3
A	1	SER	-	expression tag	UNP Q8TAF3
A	581	GLY	-	expression tag	UNP Q8TAF3
A	582	ASN	-	expression tag	UNP Q8TAF3
A	583	SER	-	expression tag	UNP Q8TAF3

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O P 5 4 1	0	0
2	A	1	Total O P 5 4 1	0	0

- Molecule 3 is GOLD ION (three-letter code: AU) (formula: Au).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	7	Total Au 7 7	0	0

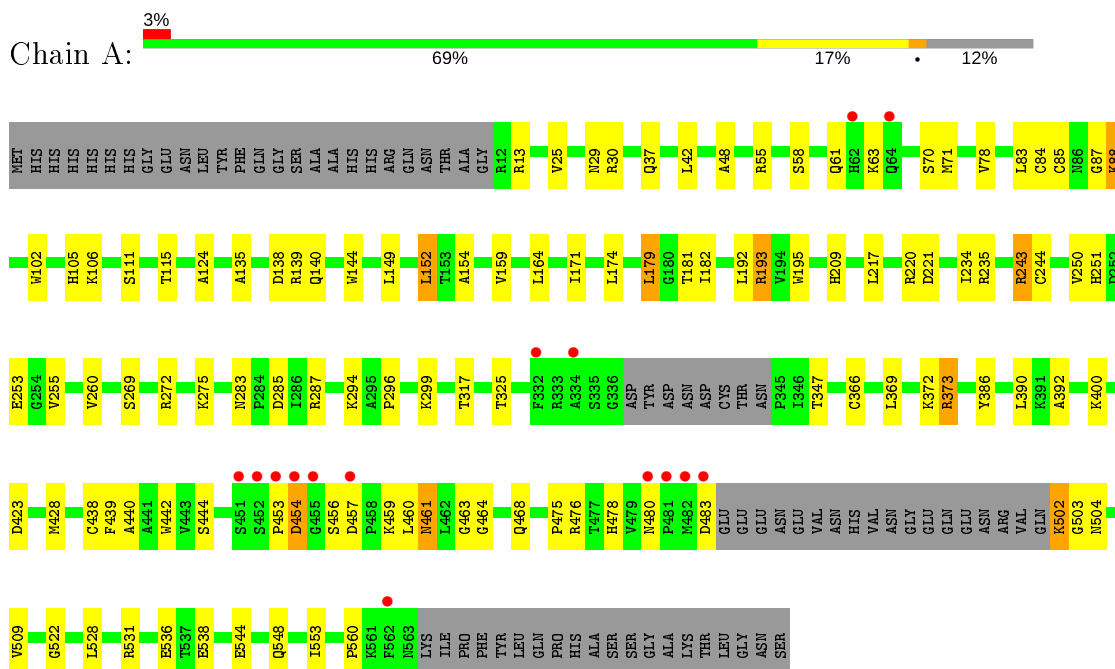
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	137	Total O 137 137	0	0

### 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: WD repeat-containing protein 48



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	139.52Å 139.52Å 235.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.33 – 3.00 49.33 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.33-3.00) 97.7 (49.33-3.00)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.36 (at 3.01Å)	Xtrriage
Refinement program	BUSTER-TNT 2.11.5	Depositor
R, $R_{free}$	0.186 , 0.233 0.187 , 0.226	Depositor DCC
$R_{free}$ test set	2000 reflections (8.47%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	66.9	Xtrriage
Anisotropy	0.282	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 73.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4292	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	69.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, AU

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	0.50	0/4224	0.75	1/5732 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	456	SER	C-N-CA	5.50	135.44	121.70

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	4138	0	4132	39	0
2	A	10	0	0	0	0
3	A	7	0	0	0	0
4	A	137	0	0	0	0
All	All	4292	0	4132	39	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 5.

All (39) 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:37:GLN:HE21	1:A:83:LEU:H	1.32	0.77
1:A:29:ASN:HA	1:A:55:ARG:HH11	1.53	0.71
1:A:149:LEU:HA	1:A:152:LEU:HD12	1.74	0.70
1:A:243:ARG:HD3	1:A:244:CYS:H	1.59	0.68
1:A:124:ALA:HB2	1:A:174:LEU:HB3	1.81	0.62
1:A:37:GLN:HG2	1:A:83:LEU:HB2	1.82	0.62
1:A:115:THR:HG21	1:A:144:TRP:CZ2	2.33	0.62
1:A:115:THR:HG21	1:A:144:TRP:HZ2	1.65	0.61
1:A:71:MET:HB3	1:A:102:TRP:CZ3	2.36	0.61
1:A:48:ALA:HB1	1:A:78:VAL:HG12	1.84	0.59
1:A:88:LYS:HG2	1:A:105:HIS:CD2	2.41	0.56
1:A:272:ARG:HA	1:A:296:PRO:HB3	1.88	0.54
1:A:243:ARG:HD3	1:A:244:CYS:N	2.23	0.54
1:A:439:PHE:HA	1:A:461:ASN:HD21	1.73	0.54
1:A:475:PRO:HA	1:A:478:HIS:HD1	1.73	0.53
1:A:502:LYS:HD3	1:A:503:GLY:H	1.74	0.52
1:A:85:CYS:O	1:A:88:LYS:HB2	2.11	0.51
1:A:179:LEU:HD21	1:A:220:ARG:HE	1.76	0.50
1:A:138:ASP:HB3	1:A:140:GLN:HG2	1.94	0.50
1:A:475:PRO:HA	1:A:478:HIS:ND1	2.28	0.49
1:A:83:LEU:HG	1:A:87:GLY:HA2	1.95	0.48
1:A:30:ARG:HH11	1:A:428:MET:HG2	1.79	0.48
1:A:55:ARG:HG2	1:A:70:SER:HB2	1.96	0.47
1:A:369:LEU:CD1	1:A:386:TYR:HE1	2.29	0.46
1:A:453:PRO:HA	1:A:454:ASP:HA	1.58	0.45
1:A:373:ARG:HH11	1:A:390:LEU:HB3	1.80	0.45
1:A:392:ALA:HB3	1:A:509:VAL:HG22	1.98	0.45
1:A:234:ILE:HD12	1:A:250:VAL:HG21	1.99	0.44
1:A:251:HIS:HE1	1:A:269:SER:OG	2.00	0.44
1:A:468:GLN:HB2	1:A:504:ASN:HB3	2.00	0.44
1:A:553:ILE:HD12	1:A:560:PRO:HD3	2.00	0.43
1:A:461:ASN:HD22	1:A:464:GLY:H	1.66	0.43
1:A:179:LEU:HB2	1:A:181:THR:HG23	2.01	0.43
1:A:135:ALA:HB1	1:A:171:ILE:HD13	2.02	0.42
1:A:42:LEU:HD11	1:A:105:HIS:CE1	2.55	0.41
1:A:193:ARG:HB2	1:A:195:TRP:CH2	2.55	0.41
1:A:209:HIS:CE1	1:A:235:ARG:HD2	2.56	0.41
1:A:283:ASN:HD21	1:A:285:ASP:HB3	1.86	0.41
1:A:438:CYS:O	1:A:463:GLY:HA3	2.21	0.41

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	520/598 (87%)	478 (92%)	38 (7%)	4 (1%)	19 57

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	154	ALA
1	A	480	ASN
1	A	440	ALA
1	A	522	GLY

### 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	460/521 (88%)	409 (89%)	51 (11%)	6 25

All (51) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	13	ARG
1	A	25	VAL
1	A	58	SER
1	A	61	GLN
1	A	63	LYS
1	A	84	CYS
1	A	88	LYS

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	106	LYS
1	A	111	SER
1	A	139	ARG
1	A	152	LEU
1	A	159	VAL
1	A	164	LEU
1	A	179	LEU
1	A	182	ILE
1	A	192	LEU
1	A	193	ARG
1	A	217	LEU
1	A	221	ASP
1	A	243	ARG
1	A	253	GLU
1	A	255	VAL
1	A	260	VAL
1	A	275	LYS
1	A	287	ARG
1	A	294	LYS
1	A	299	LYS
1	A	317	THR
1	A	325	THR
1	A	347	THR
1	A	366	CYS
1	A	372	LYS
1	A	373	ARG
1	A	400	LYS
1	A	423	ASP
1	A	442	TRP
1	A	444	SER
1	A	454	ASP
1	A	457	ASP
1	A	459	LYS
1	A	460	LEU
1	A	461	ASN
1	A	476	ARG
1	A	483	ASP
1	A	502	LYS
1	A	528	LEU
1	A	531	ARG
1	A	536	GLU
1	A	538	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	544	GLU
1	A	548	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	37	GLN
1	A	105	HIS
1	A	242	GLN
1	A	251	HIS
1	A	259	GLN
1	A	283	ASN
1	A	322	ASN
1	A	461	ASN

### 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 9 ligands modelled in this entry, 7 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	PO4	A	602	-	4,4,4	2.46	1 (25%)	6,6,6	0.49	0
2	PO4	A	601	-	4,4,4	2.50	1 (25%)	6,6,6	0.52	0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	PO4	P-O1	4.15	1.60	1.50
2	A	602	PO4	P-O1	4.09	1.60	1.50

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 [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	526/598 (87%)	-0.13	15 (2%) 51 23	39, 66, 117, 150	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	453	PRO	4.8
1	A	452	SER	4.0
1	A	481	PRO	4.0
1	A	480	ASN	3.8
1	A	483	ASP	3.7
1	A	451	SER	3.6
1	A	454	ASP	3.3
1	A	62	HIS	2.8
1	A	562	PHE	2.5
1	A	455	GLY	2.3
1	A	457	ASP	2.2
1	A	332	PHE	2.2
1	A	334	ALA	2.1
1	A	64	GLN	2.1
1	A	482	MET	2.1

### 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	PO4	A	602	5/5	0.58	0.66	105,105,106,106	5
3	AU	A	607	1/1	0.84	0.20	160,160,160,160	1
3	AU	A	606	1/1	0.92	0.15	125,125,125,125	1
2	PO4	A	601	5/5	0.94	0.22	114,115,116,118	0
3	AU	A	608	1/1	0.94	0.22	116,116,116,116	1
3	AU	A	604	1/1	0.97	0.09	145,145,145,145	1
3	AU	A	609	1/1	0.98	0.20	175,175,175,175	1
3	AU	A	605	1/1	0.98	0.10	88,88,88,88	1
3	AU	A	603	1/1	0.99	0.08	141,141,141,141	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.