



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 5, 2020 – 07:36 PM EST

PDB ID : 6J4F  
Title : Crystal structure of the AtWRKY2 domain  
Authors : Xu, Y.P.; Xu, H.; Wang, B.; Su, X.D.  
Deposited on : 2019-01-08  
Resolution : 2.40 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 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 : 2.4  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.4

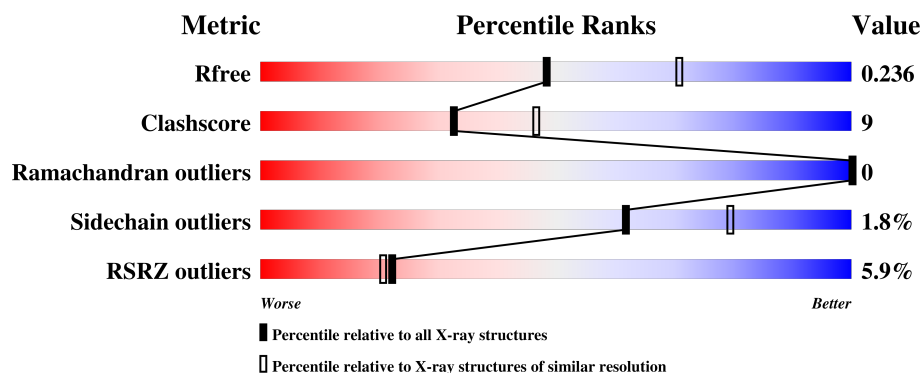
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	111664	3481 (2.40-2.40)
Clashscore	122126	3956 (2.40-2.40)
Ramachandran outliers	120053	3897 (2.40-2.40)
Sidechain outliers	120020	3898 (2.40-2.40)
RSRZ outliers	108989	3386 (2.40-2.40)

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	B	82	<div> <div>6%</div> <div> <div></div> <div>54%</div> <div>22%</div> <div>•</div> <div>23%</div> </div> </div>
1	F	82	<div> <div>7%</div> <div> <div></div> <div>59%</div> <div>18%</div> <div></div> <div>23%</div> </div> </div>
2	C	15	<div> <div></div> <div> <div>67%</div> <div>33%</div> </div> </div>
2	G	15	<div> <div></div> <div> <div>60%</div> <div>40%</div> </div> </div>
3	D	15	<div> <div></div> <div> <div>87%</div> <div>13%</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	H	15	

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
4	ZN	B	401	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2265 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 Probable WRKY transcription factor 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	63	Total	C	N	O	S	0	0	0
			517	326	97	92	2			
1	F	63	Total	C	N	O	S	0	0	0
			520	329	97	92	2			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	258	MET	-	initiating methionine	UNP Q9FG77
B	332	LEU	-	expression tag	UNP Q9FG77
B	333	GLU	-	expression tag	UNP Q9FG77
B	334	HIS	-	expression tag	UNP Q9FG77
B	335	HIS	-	expression tag	UNP Q9FG77
B	336	HIS	-	expression tag	UNP Q9FG77
B	337	HIS	-	expression tag	UNP Q9FG77
B	338	HIS	-	expression tag	UNP Q9FG77
B	339	HIS	-	expression tag	UNP Q9FG77
F	258	MET	-	initiating methionine	UNP Q9FG77
F	332	LEU	-	expression tag	UNP Q9FG77
F	333	GLU	-	expression tag	UNP Q9FG77
F	334	HIS	-	expression tag	UNP Q9FG77
F	335	HIS	-	expression tag	UNP Q9FG77
F	336	HIS	-	expression tag	UNP Q9FG77
F	337	HIS	-	expression tag	UNP Q9FG77
F	338	HIS	-	expression tag	UNP Q9FG77
F	339	HIS	-	expression tag	UNP Q9FG77

- Molecule 2 is a DNA chain called DNA (5'-D(\*AP\*GP\*CP\*CP\*TP\*TP\*TP\*GP\*AP\*CP\*CP\*AP\*GP\*CP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	15	Total	C	N	O	P	0	0	0
			303	145	56	88	14			
2	G	15	Total	C	N	O	P	0	0	0
			303	145	56	88	14			

- Molecule 3 is a DNA chain called DNA (5'-D(\*TP\*CP\*GP\*CP\*TP\*GP\*GP\*TP\*CP\*AP\*AP\*AP\*GP\*GP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	15	Total	C	N	O	P	0	0	0
			306	146	58	88	14			
3	H	15	Total	C	N	O	P	0	0	0
			306	146	58	88	14			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Zn	0	0
			1	1		
4	F	1	Total	Zn	0	0
			1	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	2	Total	O	0	0
			2	2		
5	C	1	Total	O	0	0
			1	1		
5	D	1	Total	O	0	0
			1	1		
5	F	1	Total	O	0	0
			1	1		
5	G	1	Total	O	0	0
			1	1		
5	H	2	Total	O	0	0
			2	2		

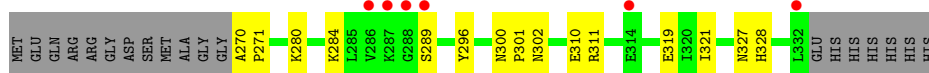
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

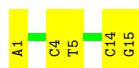
- Molecule 1: Probable WRKY transcription factor 2



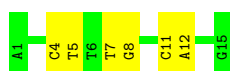
- Molecule 1: Probable WRKY transcription factor 2



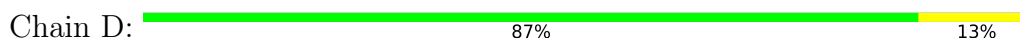
- Molecule 2: DNA (5'-D(\*AP\*GP\*CP\*CP\*TP\*TP\*TP\*GP\*AP\*CP\*CP\*AP\*GP\*CP\*G)-3')



- Molecule 2: DNA (5'-D(\*AP\*GP\*CP\*CP\*TP\*TP\*TP\*GP\*AP\*CP\*CP\*AP\*GP\*CP\*G)-3')



- Molecule 3: DNA (5'-D(\*TP\*CP\*GP\*CP\*TP\*GP\*GP\*TP\*CP\*AP\*AP\*AP\*GP\*GP\*C)-3')



- Molecule 3: DNA (5'-D(\*TP\*CP\*GP\*CP\*TP\*GP\*GP\*TP\*CP\*AP\*AP\*AP\*GP\*GP\*C)-3')

Chain H:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	38.10Å 48.30Å 69.10Å 93.70° 87.30° 97.50°	Depositor
Resolution (Å)	29.58 – 2.40 29.58 – 2.40	Depositor EDS
% Data completeness (in resolution range)	93.5 (29.58-2.40) 93.9 (29.58-2.40)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.58 (at 2.39Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.188 , 0.226 0.204 , 0.236	Depositor DCC
$R_{free}$ test set	1794 reflections (10.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.3	Xtriage
Anisotropy	0.258	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 36.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2265	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.76% 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:  
ZN

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	B	0.64	0/531	0.67	0/713
1	F	0.43	0/534	0.60	0/717
2	C	0.85	0/339	1.04	0/521
2	G	0.71	0/339	0.94	0/521
3	D	0.80	0/343	0.98	0/528
3	H	0.73	0/343	0.88	0/528
All	All	0.68	0/2429	0.84	0/3528

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 [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	B	517	0	507	17	0
1	F	520	0	515	12	0
2	C	303	0	170	3	0
2	G	303	0	170	3	0
3	D	306	0	170	1	0
3	H	306	0	170	3	0
4	B	1	0	0	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	1	0	0	0	0
5	B	2	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
5	F	1	0	0	0	0
5	G	1	0	0	0	0
5	H	2	0	0	0	0
All	All	2265	0	1702	36	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 36 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:278:TRP:HA	1:B:297:LYS:O	1.70	0.92
1:B:298:CYS:HG	4:B:401:ZN:ZN	0.77	0.91
3:H:3:DG:H2'	3:H:4:DC:C6	2.17	0.80
1:B:290:GLU:HB2	1:B:313:ARG:HE	1.46	0.79
1:B:281:TYR:HE1	1:B:296:TYR:O	1.76	0.68

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	B	61/82 (74%)	59 (97%)	2 (3%)	0	100	100
1	F	61/82 (74%)	58 (95%)	3 (5%)	0	100	100
All	All	122/164 (74%)	117 (96%)	5 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	B	54/70 (77%)	52 (96%)	2 (4%)	37	56
1	F	55/70 (79%)	55 (100%)	0	100	100
All	All	109/140 (78%)	107 (98%)	2 (2%)	62	79

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	293	ARG
1	B	297	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	283	GLN
1	F	300	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

Of 2 ligands modelled in this entry, 2 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

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	B	63/82 (76%)	0.60	5 (7%) 12 11	34, 50, 82, 99	0
1	F	63/82 (76%)	0.67	6 (9%) 8 7	38, 59, 85, 90	0
2	C	15/15 (100%)	0.04	0 100 100	47, 55, 82, 82	0
2	G	15/15 (100%)	0.02	0 100 100	57, 74, 97, 98	0
3	D	15/15 (100%)	-0.06	0 100 100	43, 65, 87, 88	0
3	H	15/15 (100%)	0.15	0 100 100	60, 74, 107, 107	0
All	All	186/224 (83%)	0.44	11 (5%) 22 21	34, 60, 90, 107	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	332	LEU	4.7
1	F	314	GLU	2.9
1	F	289	SER	2.6
1	F	288	GLY	2.5
1	F	287	LYS	2.5

### 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](#)

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
4	ZN	B	401	1/1	0.96	0.13	44,44,44,44	0
4	ZN	F	401	1/1	1.00	0.12	52,52,52,52	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.