



## Full wwPDB X-ray Structure Validation Report ⓘ

May 17, 2020 – 08:01 pm BST

PDB ID : 6QTR  
Title : Crystal structure of a mutant Arabidopsis WD40 domain in complex with a transcription factor  
Authors : Hothorn, M.; Lau, K.  
Deposited on : 2019-02-25  
Resolution : 1.37 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



## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 5519 atoms, of which 2599 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 E3 ubiquitin-protein ligase COP1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	307	5026	1592	2498	430	484	22	0	20	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	346	GLY	-	expression tag	UNP P43254
A	347	ALA	-	expression tag	UNP P43254
A	348	MET	-	expression tag	UNP P43254

- Molecule 2 is a protein called Transcription factor HY5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
2	B	11	181	60	88	17	16	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	38	ACE	-	acetylation	UNP O24646
B	49	TYR	GLU	conflict	UNP O24646

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	1
			25	6	13	6		

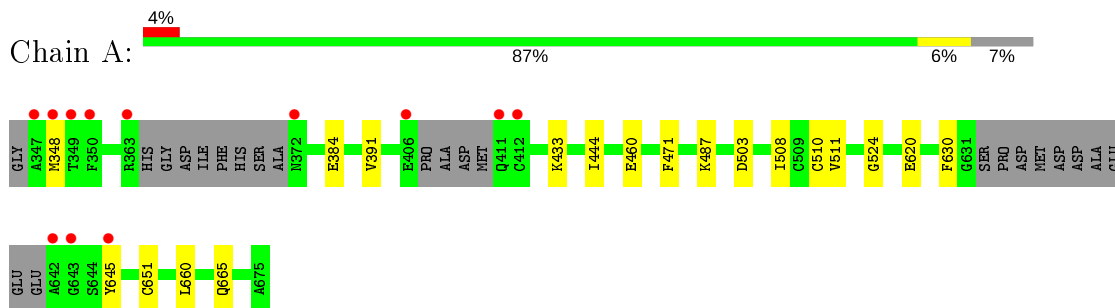
- Molecule 5 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	A	269	Total 271	O 271	0	2
5	B	6	Total 6	O 6	0	0

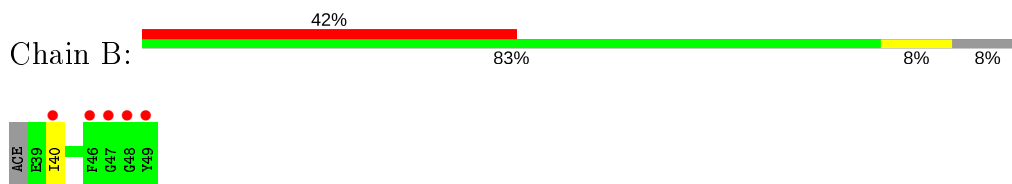
### 3 Residue-property plots [i](#)

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: E3 ubiquitin-protein ligase COP1



- Molecule 2: Transcription factor HY5



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.26Å 55.06Å 103.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.74 – 1.37 48.61 – 1.14	Depositor EDS
% Data completeness (in resolution range)	97.4 (43.74-1.37) 84.0 (48.61-1.14)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.46 (at 1.14Å)	Xtrriage
Refinement program	PHENIX 1.13_2998	Depositor
R, $R_{free}$	0.120 , 0.161 0.120 , 0.161	Depositor DCC
$R_{free}$ test set	2853 reflections (3.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	9.1	Xtrriage
Anisotropy	0.067	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 45.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	5519	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.40% 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: GOL, CSO, CME, SO4

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	0.42	0/2582	0.68	0/3485
2	B	0.37	0/95	0.46	0/126
All	All	0.42	0/2677	0.67	0/3611

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	A	2528	2498	2506	15	0
2	B	93	88	88	1	0
3	A	10	0	0	0	0
4	A	12	13	16	0	0
5	A	271	0	0	6	3
5	B	6	0	0	0	1
All	All	2920	2599	2610	15	3

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 3.

All (15) 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:510[C]:CSO:OD	5:A:801:HOH:O	2.12	0.67
1:A:444[A]:ILE:HD12	1:A:460:GLU:HG2	1.80	0.63
1:A:510[A]:CSO:SG	1:A:524:GLY:HA3	2.42	0.59
1:A:511[A]:VAL:CG2	5:A:894:HOH:O	2.54	0.55
1:A:348:MET:HE1	1:A:620:GLU:HG3	1.90	0.53
1:A:384:GLU:OE1	1:A:433:LYS:NZ	2.40	0.53
1:A:444[A]:ILE:HD13	5:A:931:HOH:O	2.09	0.51
1:A:511[A]:VAL:HG23	5:A:894:HOH:O	2.09	0.51
1:A:651[B]:CYS:SG	1:A:660:LEU:HD12	2.51	0.50
1:A:508:ILE:HD11	5:A:813:HOH:O	2.15	0.47
1:A:645:TYR:CD1	1:A:665:GLN:HG2	2.50	0.47
1:A:508:ILE:CD1	5:A:813:HOH:O	2.62	0.46
1:A:391[A]:VAL:HG11	2:B:40:ILE:HG21	1.99	0.45
1:A:487:LYS:HD2	1:A:503:ASP:OD1	2.16	0.45
1:A:645:TYR:CE1	1:A:665:GLN:HG2	2.55	0.41

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:841:HOH:O	5:A:1018:HOH:O[4_555]	2.04	0.16
5:A:976:HOH:O	5:A:1062:HOH:O[3_455]	2.12	0.08
5:A:1021:HOH:O	5:B:106:HOH:O[3_445]	2.14	0.06

## 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	313/330 (95%)	302 (96%)	11 (4%)	0	100	100
2	B	9/12 (75%)	9 (100%)	0	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	322/342 (94%)	311 (97%)	11 (3%)	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	A	289/290 (100%)	287 (99%)	2 (1%)	84	65
2	B	9/9 (100%)	9 (100%)	0	100	100
All	All	298/299 (100%)	296 (99%)	2 (1%)	84	65

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

Mol	Chain	Res	Type
1	A	471	PHE
1	A	630	PHE

Some 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](#)

7 non-standard protein/DNA/RNA residues are modelled in this entry.

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	$\# Z  > 2$
1	CSO	A	425[B]	-	3,6,7	0.73	0	0,6,8	0.00	-
1	CSO	A	510[B]	-	3,6,7	0.56	0	0,6,8	0.00	-
1	CSO	A	425[A]	-	3,6,7	0.74	0	0,6,8	0.00	-
1	CSO	A	510[C]	-	3,6,7	0.51	0	0,6,8	0.00	-
1	CME	A	394	1	8,9,10	0.93	0	5,9,11	0.82	0
1	CSO	A	492	1	3,6,7	0.68	0	0,6,8	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CSO	A	425[B]	-	-	0/1/5/7	-
1	CSO	A	510[B]	-	-	0/1/5/7	-
1	CSO	A	425[A]	-	-	0/1/5/7	-
1	CSO	A	510[C]	-	-	0/1/5/7	-
1	CME	A	394	1	-	0/5/8/10	-
1	CSO	A	492	1	-	0/1/5/7	-

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	510[C]	CSO	1	0

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

4 ligands are modelled in this entry.

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
3	SO4	A	702	-	4,4,4	0.14	0	6,6,6	0.05	0
3	SO4	A	701	-	4,4,4	0.15	0	6,6,6	0.40	0
4	GOL	A	703[A]	-	5,5,5	0.99	0	5,5,5	0.82	0
4	GOL	A	703[B]	-	5,5,5	1.02	0	5,5,5	0.99	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	A	703[A]	-	-	0/4/4/4	-
4	GOL	A	703[B]	-	-	0/4/4/4	-

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 [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	303/330 (91%)	-0.31	12 (3%) 38 40	6, 10, 30, 88	0
2	B	11/12 (91%)	1.95	5 (45%) 0 0	11, 37, 63, 75	0
All	All	314/342 (91%)	-0.23	17 (5%) 25 26	6, 10, 39, 88	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	40	ILE	6.7
2	B	49	TYR	6.6
1	A	350	PHE	5.9
1	A	348	MET	5.8
1	A	347	ALA	5.7
1	A	642	ALA	5.3
1	A	643	GLY	5.1
1	A	363	ARG	4.4
1	A	412	CYS	3.1
2	B	46	PHE	2.8
1	A	406	GLU	2.6
1	A	349	THR	2.5
2	B	48	GLY	2.4
2	B	47	GLY	2.2
1	A	411	GLN	2.1
1	A	645	TYR	2.1
1	A	372	ASN	2.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [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( $\text{\AA}^2$ )	Q<0.9
1	CSO	A	510[B]	7/8	0.98	0.09	6,8,11,12	8
1	CSO	A	510[C]	7/8	0.98	0.09	6,7,15,18	8
1	CSO	A	510[A]	6/8	0.98	0.09	6,7,7,7	6
1	CME	A	394	10/11	0.98	0.07	9,15,30,33	0
1	CSO	A	425[A]	7/8	0.99	0.07	6,8,11,13	8
1	CSO	A	425[B]	7/8	0.99	0.07	6,7,10,12	8
1	CSO	A	492	7/8	0.99	0.06	8,10,16,19	0

### 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( $\text{\AA}^2$ )	Q<0.9
3	SO4	A	702	5/5	0.94	0.10	45,45,46,47	0
3	SO4	A	701	5/5	0.98	0.08	14,17,18,19	5
4	GOL	A	703[A]	6/6	0.98	0.07	8,11,14,16	13
4	GOL	A	703[B]	6/6	0.98	0.07	14,18,21,21	12

### 6.5 Other polymers [i](#)

There are no such residues in this entry.